(12)

DECISION SCIENCE CONSORTIUM, INC.

ALTERNATE THEORIES OF BELIEF AND THE IMPLICATIONS FOR INCOHERENCE, RECONCILIATION, AND SENSITIVITY ANALYSIS

Anthony N. S. Freeling

Decision Science Consortium, Inc.
7700 Leesburg Pike, Suite 421
Falls Church, Virginia 22043

November 1981

Prepared for:

OFFICE OF NAVAL RESEARCH
Mathematical Sciences Division
Naval Analysis Programs
800 N. Quincy Street
Arlington, Virginia 22217

Technical Report No. 81-4

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	
2. GOVT ACCESSION NO.	AD-A113 193
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)	ALTERNATE THEORIES OF BELIEF AND THE IMPLICATIONS FOR INCOHERENCE, RECONCILIATION, AND SENSITIVITY ANALYSIS
5. TYPE OF REPORT & PERIOD COVERED	Technical
6. PERFORMING ORG. REPORT NUMBER	81-4
7. AUTHOR(s)	Anthony N. S. Freeling
8. CONTRACT OR GRANT NUMBER(s)	N00014-81-C-0330
9. PERFORMING ORGANIZATION NAME AND ADDRESS	DECISION SCIENCE CONSORTIUM, Inc. 7700 Leesburg Pike, Suite 421 Falls Church, Virginia 22043
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	61153N RR 014-11-01 NR 277-271
11. CONTROLLING OFFICE NAME AND ADDRESS	Naval Analysis Prog Mathematical and Information Sciences Division OFFICE OF NAVAL RESEARCH, Code 431 800 N. Quincy St., Arlington, VA 22217
12. REPORT DATE	November 1981
13. NUMBER OF PAGES	84
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	
15. SECURITY CLASS. (of this report)	Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

16. DISTRIBUTION STATEMENT (of this Report)

Approved for publication release: distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Decision Theory
Belief Functions
Incoherence
Sensitivity Analysis
Upper and Lower Probabilities
Fuzzy Probabilities

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we examine various theories of belief alternate to subjective probability. We examine the motivation for each such theory, and place each theory within the context of decision aiding.

Initially, we examine the role of normative theories of decision making and belief, distinguishing carefully between the terms normative and prescriptive. We conclude that the decision analysis paradigm is compelling normatively,

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20.
Abstract (continued)

but not prescriptively. We ~~then~~ discuss inconsistency with the decision analysis axioms, and define incoherence as the potential for forming inconsistent judgments. We propose that decision analysis is a means for reducing incoherence. We further argue that sensitivity analysis is used as a means for countering incoherence, and that many extended theories of belief may be viewed as formal justifications for sensitivity analysis.

We ~~then~~ examine theories of upper and lower probabilities from this perspective, together with second-order and fuzzy probabilities. We point out in each case the similarities with the other theories, and look at the problem of eliciting the relevant information from a decision maker.

We examine in detail the theory of Belief Functions of Shafer (1976), one form of upper and lower probability. We discuss this in a theory of evidence, rather than of belief, and show how such a theory might provide advantages over traditional Bayesian methods. We conclude, however, that the assessment problem has not been solved.

We then look at various measures of belief which have only ordinal properties; including inductive probabilities (Cohen, 1977) and possibility theory arising from fuzzy sets. We show that these too are theories of evidence, but with greater potential for application due to reduced assessment difficulties.

Finally, we look again at the implication of our work for practical decision analysis and sensitivity analysis. We conclude that the "divide and conquer" strategy is unsatisfactory when a sensitivity analysis is considered, since some of the relevant information from a decision maker is lost. We stress that to use the maximal information, the entire belief structure should be modeled, and we make tentative suggestions towards developing a new methodology based on these observations.

Our overall conclusion is that the present-day practice of decision analysis is adequate, but that it might be refined, and sensitivity analysis improved, if note were taken of these alternate theories of belief.

ALTERNATE THEORIES OF BELIEF AND THE IMPLICATIONS FOR INCOHERENCE, RECONCILIATION, AND SENSITIVITY ANALYSIS

By

Anthony N. S. Freeling

Prepared for:

Office of Naval Research
Mathematical Sciences Division
Naval Analysis Programs
Contract No. N00014-81-C-0330

November 1981

Decision Science Consortium, Inc.
7700 Leesburg Pike, Suite 421
Falls Church, Virginia 22043
(703) 790-0510

SUMMARY

In this paper we examine various theories of belief alternate to subjective probability. We examine the motivation for each such theory, and place each theory within the context of decision aiding.

Initially, we examine the role of normative theories of decision making and belief, distinguishing carefully between the terms normative and prescriptive. We conclude that the decision analysis paradigm is compelling normatively, but not prescriptively. We then discuss inconsistency with the decision analysis axioms, and define incoherence as the potential for forming inconsistent judgments. We propose that decision analysis is a means for reducing incoherence. We further argue that sensitivity analysis is used as a means for countering incoherence, and that many extended theories of belief may be viewed as formal justifications for sensitivity analysis.

We then examine theories of upper and lower probabilities from this perspective, together with second-order and fuzzy probabilities. We point out in each case the similarities with the other theories, and look at the problem of eliciting the relevant information from a decision maker.

We examine in detail the theory of Belief Functions of Shafer (1976), one form of upper and lower probability. We discuss this in a theory of evidence, rather than of belief, and show how such a theory might provide advantages over traditional Bayesian methods. We conclude, however, that the assessment problem has not been solved.

We then look at various measures of belief which have only ordinal properties; including inductive probabilities (Cohen, 1977) and possibility theory arising from fuzzy sets. We show that these too are theories of evidence, but with greater potential for application due to reduced assessment difficulties.

Finally, we look again at the implication of our work for practical decision analysis and sensitivity analysis. We conclude that the "divide and conquer" strategy is unsatisfactory when a sensitivity analysis is considered, since some of the relevant information from a decision maker is lost. We stress that to use the maximal information, the entire belief structure should be modeled, and we make tentative suggestions towards developing a new methodology based on these observations.

Our overall conclusion is that the present-day practice of decision analysis is adequate, but that it might be refined, and sensitivity analysis improved, if note were taken of these alternate theories of belief.

ACKNOWLEDGMENTS

This work was funded by the Office of Naval Research, Naval Analysis Program under Contract No. N00014-81-C-0330.

TABLE OF CONTENTS

1.0 INTRODUCTION

The work reported in this paper has arisen out of previous work on Reconciling Incoherent Judgment (RIJ). That work addressed the problem faced by an analyst when a decision maker (DM) provides subjective assessments which fail to satisfy the axioms of probability and/or utility theory. The DM is then being inconsistent, or exhibiting incoherence. In our previous work on RIJ (Brown and Lindley, 1981; Lindley, Tversky, and Brown, 1979; Freeling, 1980b, 1981a, b) we have examined the possibility of producing mathematical techniques to provide a single set of consistent, reconciled values from the inconsistent set provided by the DM. While we have had some success in producing such techniques, it has become apparent that there is no unique reconciliation of a set of inconsistent values. We have also found that it is necessary to ask further questions of the DM in order to discover more about his/her belief structure. Such further questions may concern the precision of the originally assessed values, or the DM's confidence in those values, or the amount of information captured by each assessment. In each case, these higher-order assessments appear to be necessary because the ordinary decision-analytic procedure of assessing probabilities and utilities in a decision tree has failed to model adequately the whole of the DM's belief structure.

The starting point for the current work has been that since "classical" subjective probability theory has failed to model the situation adequately, we should look at certain other mathematical theories of belief that have been developed. If such a theory were sufficiently rich, it may be that the inconsistency discovered relative to the probability calculus would be acceptable under the alternative calculus. Failing that, the perspective offered by such a theory might provide insights into improved ways of performing a reconciliation.

We therefore, looked in detail at work that has been performed on theories of belief alternative to subjective probability. In particular we looked at axiom systems producing upper and lower probabilities (Koopman, 1940a, b; Good, 1962; Smith, 1961; Dempster, 1967; Suppes, 1974; Nau, 1981); at the theory of belief functions (Shafer, 1976); at the use of hierarchical probability structures (Good, 1952; Lindley, Tversky, and Brown, 1979); at various uses of fuzzy set theory (Zadeh, 1965, 1978; Watson, Weiss, and Donnell, 1979; Yager, 1979; Freeling, 1979, 1980a, c, d); and at some related work of L.J. Cohen (1973, 1977, 1979, 1980) and of Shackle (1969).

Each of these theories has been well-developed in an abstract form. Each theory weakens, in some way, the strength of the axioms that lead to belief being measured by probabilities on what is, essentially, a ratio scale. This is done by allowing vagueness in assessments, to produce ranges of values on the ratio scale, producing upper and lower probabilities; or by producing a scale which has only ordinal properties. We also found that certain of the theories based the modeling of belief

on the concept of chance (viz. subjective probability, upper and lower probabilities, hierarchical probabilities), whereas some are best interpreted in terms of weights of evidence (viz. belief functions, fuzzy set theory, Cohen's inductive probabilities).

As our study of the mathematics and underlying philosophy of each of these theories proceeded, we were forced to re-examine our motivation for the study, and to rethink our views on incoherence, reconciliation, and the aim of decision analysis. This led us to the conclusion that the sensitivity analysis is a vital part of any decision analysis, to a greater extent than is usually acknowledged. We present our reasoning behind this conclusion in Section 2.0. In the three sections after that we discuss the various alternate theories of belief in detail. We have concentrated on their possible practical use in decision aiding or inference, by concentrating on the behavioral assumptions implicit in their foundations. We discuss the strengths and weaknesses of each one, and look at the links and differences among them. In particular, we look at how these theories may provide axiomatic justification for sensitivity analysis; and what guidance can be given for the performance of the analysis.

In Section 6.0 we summarize our results and present our conclusions. The current research has been of a divergent nature--we have looked at a wide range of literature and attempted to place it within a common context; we have examined foundations of our practice and attempted to generate a coherent philosophical basis; but we have not developed in detail any specific procedures. Such research will, we hope, be

embarked upon soon. The present paper will have achieved its aim if it causes practicing decision analysts to re-examine the philosophy behind their work, to be aware of the parallel but distinct theories we discuss here, and stimulates further work on performance of sensitivity analyses.

2.0 BACKGROUND

The standard decision-analytic paradigm assumes that decision makers (DMs) are capable of quantifying their uncertainties and values in the form of probabilities and utilities, respectively (Raiffa, 1968; Edwards, 1954). It is further assumed that these quantities may be assessed to an arbitrary degree of precision. In practice, such an assumption has been found to be false. Much psychological and practical work has shown that decision makers may consistently violate the axioms that lead to numerical scales for probabilities and utilities. Similarly, the RIJ studies have been developed out of the observation that DMs will often produce values that are inconsistent with the laws of probability. They will also often protest that they have very little confidence in a certain assessment; that they do not wish to be committed to any particular value. Such findings should not be viewed with surprise--they may each be interpreted in terms of the limited ability of human beings to handle and process information (Slovic, 1972).

In terms of the practical application of decision analysis as a decision aid, these problems have usually been dealt with by carrying out sensitivity analyses at the end of an analysis, in order to see whether shifting the assessed values produces a shift in preferred alternatives. By studying the results of such sensitivity analyses far greater insight into the nature of the problem can be obtained than from the basic results.

However, the standard axiom systems do not allow for the possibility that a sensitivity analysis might be necessary. The argument (see, e.g., Savage, 1954) is that, since a subjective probability is given by the DM, that _is_ the subjective probability for that DM, and there is no meaning to producing arguments of the form "what if the probability were in fact slightly different." There has thus been a good deal of work performed that is aimed at producing systems of belief that differ from standard probability theory by allowing ranges of probability, rather than the point values given in the standard paradigm. These ranges could then be understood as the range over which the sensitivity analysis should be performed.

2.1 Upper and Lower Probabilities

The different theories that have been produced and which we shall discuss are presented in a variety of different ways. We shall see, however, that they may in fact be viewed as falling into one of just two categories. The first category is that of producing _upper_ and _lower_ probabilities. The second category consists of what we shall term _ordinal_ measures.

The basic concept of upper and lower probabilities is very simple, and is directly related to our discussion of ranges of probability rather than point estimates. For a particular event A, say, the lower probability is simply the lower bound on the possible range of probabilities, and the upper probability is the upper bound. Thus, if the lower probability is equal to the upper one, this value is

the only possible value in the range, which is, therefore, an ordinary (point) probability. So we see that a theory of belief which is based on upper and lower probabilities contains ordinary probability theory as a subset. We shall term the entities which are expressed in terms of lower and upper probabilities imprecise probabilities.

Although the concept of lower and upper probabilities appears quite simple, there are several different variations on the theme. These have been derived from differing axiom systems, and have been developed with different aims in mind. We wish to examine them with regard to their potential value as part of a decision aid. As we shall discuss in the remainder of this section, the value is primarily related to the implications for the performance of sensitivity analysis.

The term "ordinal measures" we use to refer to those theories of belief with degrees of belief upon which only the mathematical operations of maximum and minimum are permissible. Such scales, therefore, require only ordinal properties. Since the concept of chance, which is the basis of probability, has stronger properties, it appears that chance is not the basis of ordinal measures. Rather, these are theories of belief based on the concept of weights of evidence. As we discuss later, this may be of value in theories of inference, but appears to be of limited value for a theory of choice.

Indeed, although our original aim was to develop an axiomatic theory of choice based upon one of these extended theories of belief, with the aim of improving upon standard DA, we came to the conclusion that this

was a vain hope. To help understand our reasons for this, we now give a personal view of the philosophy underlying the whole concept of decision analysis.

2.2 Normative versus Prescriptive

It is often stated that decision analysis is a normative theory rather than a descriptive theory of decision making (Edwards and Tversky, 1967). It has also been claimed that decision analysis is prescriptive (Raiffa, 1968). A distinction can be made between the two concepts of normative and prescriptive. A normative theory is one which outlines how a DM of unlimited intellectual ability would act if he/she held certain beliefs and values. A prescriptive theory is one which prescribes how a DM should be advised to act, once certain required information has been elicited from him/her. The distinction is that between an idealized human on the one hand, and a real, fallible one on the other. Keeney and Raiffa (1976) make this distinction in their preface.

We believe that decision analysis is correctly termed a normative theory, in that axioms leading to subjective probability (such as Savage's) and to von Neumann-Morgenstern utility deal exclusively with a perfectly rational being. Those of us who are in the business of aiding decision-making are, however, in need of a prescriptive theory. It would certainly not be considered very helpful by a DM if a decision analyst were to say that if the DM were only more rational, the indicated decision would be X, but unfortunately, due to the DM's irrationality, the analyst has no idea what should be done.

Keeney and Raiffa (1976) claim that decision analysis is prescriptive, in that an analyst is trying to aid a real DM. However, we are trying to make the distinction here between the underlying foundations of decision analysis, which we believe to be normative, and the practical application of decision analysis, which must perforce be prescriptive. Any practical analyst will have to add some heuristics to the basic theory in order to apply the techniques. We should emphasize that we do not wish normative to be seen as synonymous with objective. Two rational beings could have totally different beliefs, and provide very different probabilities and utilities, while each conforming precisely to the normative axioms.

It is in recognition of the normative nature of decision analysis and of the requirement for prescription that one embarks upon sensitivity analyses. One is, in effect, developing a pragmatic theory, saying:

a) The axioms show us how a perfect being would act - Normative Assumption

b) The DM is not perfect, so let us assume he/she deviates only slightly from the axioms, and hope that the sensitivity analysis includes somewhere his/her actual behavior - Pragmatic Assumption.

Taking this perspective it appears that the search for a set of axioms leading to a prescriptive theory must perforce be doomed to failure. A set of axioms (together with the laws of logic) <u>define</u> rationality in a given context, and on the pragmatic assumption that real DM's will not be totally rational (in this sense), the theories will be inadequate for prescription. One is in fact seeking some combination of <u>descriptive</u> theory and normative theory. A decision analyst may be correctly viewed as being on the borderline between philosophy and psychology. Put simplistically, a philosopher may say what one ought to do (normative) and a psychologist, what one does do (descriptive).

We thus wish to argue that it would be an enterprise almost certainly doomed to failure, were one to seek a fixed set of axioms to serve as the basis for a prescriptive theory of decision aiding. A DM could always violate such axioms. We do not wish, therefore, to use the theories of belief we present in following sections as the foundation for a new form of decision analysis. Rather, we view them as theories which explore the consequences of deviating from "classical" subjective probability, and which may, therefore, shed light upon the ways in which we might deviate from the standard DA paradigm in order to improve our decision aiding. Rather than an entirely new theory, we are attempting to justify and explore methods for conducting sensitivity analyses from the DA method.

2.3 Incoherence versus Inconsistency

The work reported in this paper has arisen out of our previous work on the reconciliation of inconsistent judgments (RIJ). As discussed in Section 1.0, our initial hope was to develop a theory of belief which was able to explain within its extended scope those assessments which appeared to be inconsistent in the context of the traditional theory. As intimated in Section 2.2, we now feel such an aim to have been misguided. To further understand our reasons for this we now look more deeply at the two concepts of inconsistency and incoherence.

The two words have, in the previous work on RIJ, tended to be used interchangeably. That has been, however, a mistake, and one which we feel may have led to an obfuscation of some of the important issues.

Inconsistency we shall use to refer to a set of judgments, or of assessed values, which exhibit a disagreement with the axiom system being used. (In general, in the context of the present paper, this will be the theory of subjective probability.) For example, if Pr(A) is given as 0.4, Pr(B/A) as 0.5, then if $Pr(A \vee B)$ is not 0.2, we have an inconsistency. The important point to note is that inconsistency is empirically defined--we simply check to see whether a set of judgments violates the relevant calculus. It is also apparent that inconsistency is only defined relative to the calculus under consideration. Incoherence, on the other hand, is less easily defined. There is an implication that an incoherent DM must, in some fundamental way, be acting irrationally. There are, however, obvious problems in attempting to produce an "absolute" definition of rationality or coherence. We shall be forced to retain a non-concrete definition of incoherence (although we feel that it is important for philosophers and practitioners to think carefully about the implications for definitions of rationality and coherence of the current work and other work in this area). We shall define:

> A DM is incoherent if he/she has failed correctly to integrate all the information he/she obtained with his/her belief structure.

This rather begs the question by failing to define "correctly." However, in the present context of decision analysis, we shall use the classical paradigm of DA as our reference standard. So, in this context,

> A DM is being incoherent if the potential exists in his/her belief structure for inconsistent judgments.

When using this definition, we feel that pejorative connotations of the word incoherent should be downplayed. One could easily argue that all DMs are to some extent incoherent in that various psychological work (Kahneman and Tversky, 1974, 1979) has shown one can usually "fool" subjects into inconsistent estimates.

We hypothesize that this is because coherence does not exist. At least, we can never be sure that we have explored the entire belief structure of the DM, so the potential for inconsistency will always remain. However, as the DM integrates more information into his/her belief structure, we may state that the degree of incoherence is reduced. We may now restate in the present vocabulary our view of practical DA:

> Decision Analysis aims to reduce the incoherence of a decision maker.

In other words, by eliciting probabilities and values from a decision maker, and by pointing out inconsistencies in these assessments, a decision analyst may help a DM explore his/her (incoherent) belief structure, and to change that structure in such a way as to eliminate those inconsistencies, and to reduce the potential for further such inconsistencies.

A similar argument is put forward by French (1979a, b). He argues that the role of a decision analysis is to set up a model decision maker (m.d.m.) who is like the DM, but idealized to conform to the DA

paradigm. By observing the implications of the DM's assessed values for the preferences of the m.d.m., the DM can educate him/herself and change his/her inconsistent preference-belief structure. In other words, we do not tell a DM what he/she thinks, but rather we are helping him/her to think.

We believe that this observation regarding the aim of a decision analysis has been insufficiently understood previously. This view of incoherence has been ignored in much previous work on incoherence, and also by other research on DA, both methodological and psychological.

2.4 Summary and Conclusions

Our position may be stated succinctly as follows:

1. An axiomatic theory of decision making can at best be normative, rather than prescriptive.
2. Any definition of optimal decision making, derived together with a methodological basis, will have an axiomatic theory as its foundation, and all work on inconsistency is only relative to that theory.
3. An extended theory of belief might still be violated by DMs to produce inconsistency.
4. Inconsistency cannot be dealt with within the axiomatic system that has been violated.
5. Sensitivity analysis is the activity that permits a normative theory of decision making to become a prescriptive decision aid.

With this perspective, inconsistency is not necessarily pejorative to the DM--the theory is as likely to be "at fault" as the DM. One important aspect of the axiomatic theories of belief to be discussed, is that each emphasizes that the entire belief structure must have been considered in order to assign (imprecise) probabilities to a given event. The concept from RIJ of "over-specification" is thus not over-specifying the important events (which is the entire universe for all events are relevant), but an attempt to explore further the belief structure to obtain assessments more in line with axiomatic foundations. Therefore, a reconciliation of such "over-specification" should not be an end in itself, but rather an aid to exploring the entire belief structure.

We accept the decision-analytic paradigm as correct normatively. We do not wish to attempt to displace it from its position (although at the beginning of the work reported here that had been our aim). We believe, however, coherence to be an ideal unattainable by real DMs, due to their having only finite information handling capability. The study of the extended theories of belief discussed in this paper, the study of RIJ, and the practice of sensitivity analysis all should be aimed at helping the decision analyst cope with the practical problem of incoherence of DMs. The importance of the theoretical work lies in the insights we can gain into how deviations from the norm of DA affect the conclusions.

It should thus become clear that when practicing DA, this perception of the fundamental incoherence of the DM should be borne in mind from the start of the analysis, rather than just brought in at the end in the form of a (often incomplete) sensitivity analysis. Throughout an analysis we attempt to help the DM explore his/her entire belief structure to improve his/her own coherence, and thus to approach a fuller understanding of his/her preferences between options.

3.0 UPPER AND LOWER PROBABILITIES

The basic definition of lower and upper probabilities provided in the previous section left an important question open--with what calculus are these quantities to be combined? A major strength of the theories of subjective probability is that each such theory permits the calculus of probability theory to be used. This means that the basic strategy of decision analysis, "divide-and-conquer," may be used. With this strategy probabilities which would be difficult to assess directly are split into several constituent probabilities. These may then be assessed more simply. The resulting values are then combined in a logically rigorous way to produce a value for the composite probability. The important point is that, together with an interpretation of the meanings of assessed probabilities, there are also appropriate rules for combining these probabilities.

For an extended theory of belief to be of practical value, it is necessary that a comparable calculus be available. It is in an attempt to provide such a calculus that the differing axiom systems have been developed. The structure that is imposed on the theory of belief by the axioms dictates the appropriate rules of combination. Each of the axiom systems may be viewed as a behavioral <u>explanation</u> for the inconsistency that is apparent in normal probability assessments, and which leads to the necessity for a range of probabilities. The

difficulty with developing rules for combination lies in the question of "second-order interaction." By this we refer to the phenomenon that two imprecise probabilities may have imprecisions which are closely related, e.g., we may be very unsure of the probability of events A and B, yet know that _if_ the probability of A were high, that of B would be low (an extreme example of this would occur if B were ~A). This would then put a constraint on the imprecise probability of A∨B, arising from second-order considerations. As we shall see, the way in which this interaction is modeled (or not modeled) provides a means to distinguish between the different systems producing imprecise probabilities.

In this section we examine in detail the differing forms of vague probabilities that have been proposed in the literature. In Section 6.0 we shall examine the extent to which these theories can help solve our problem. Some previous work on looking at the formal similarities and overlap between various theories of belief has been performed by Prade (1978). Although he discusses some of these similarities and differences in terms of semantic implications of the different names for their theories, his work is primarily abstract in nature. We shall attempt to conduct our examination on a more applied level.

3.1 Ranges of Probability

The most natural and simple way to extend the basic theory of probability is to relax the assumption which forms part of each axiomatic system, that humans are able to rank-order any events of different likelihood. This assumption also implies that a decision maker will

feel equally sure about each possible comparison or probability judgment. By a suitable relaxation of the system we should be able to produce vague probabilities, which we define to be ranges for each probability such that the upper and lower probabilities each conform to the classical probability calculus.

Indeed, such vague probabilities have already been implicity assumed by decision analysts when conducting sensitivity analyses. For typically in that stage of a DA the behavior of the model is observed under the supposition that probabilities were either higher or lower than actually assessed. By observing the behavior of the solution between the ranges of possible probabilities, the analyst obtains insight into the problem solution. Note that the low and high values are each operated with as if they were probabilities. To illustrate this, suppose that the probability of A were described by the range [0.2, 0.4]; and the probability of B by [0.3, 0.6]. Then supposing A and B to be independent, the probability of $A \wedge B$ could be deduced by looking at the upper and lower probabilities separately. So the lower probability would be $0.2 \times 0.3 = 0.06$, and the upper probability would be $0.6 \times 0.4 = 0.24$. Similarly, $A \vee B$ would have lower probability $(0.2 + 0.3 - 0.06) = 0.44$ and upper probability $(0.6 + 0.4 - 0.24) = 0.76$. This is the simplest extension of classical probability theory. Axiomatizing such a system of vague probabilities would thus provide a justification for the usual form of sensitivity analysis.

3.2 Previous Work

Systems of this form have been proposed by several different authors. The

idea apparently originated with Boole (1854). One of the earliest set of axioms was developed by Koopman (1940a, b). Good (1962) has provided a set of axioms which he believes to be a simplification of those of Koopman. Smith (1961, 1965) has derived upper and lower probabilities from a consideration of betting odds. Essentially, he uses the betting paradigm of subjective probability, but releases the restriction that one must bet on an event at the same odds at which one wishes to bet against that event. He argues that the restriction provides a "more definite expression of opinion" than he would wish.

The work by Smith has been extended and presented in less opaque form very recently by Nau (1981). Nau discusses the classical theory of subjective probability as proposed by De Finetti (1974). He examines the betting paradigm from the perspective of linear programming (LP). The essence of this approach is to show that the existence of a subjective probability distribution is equivalent to the existence of a set of "fair" betting prices, and to discover these prices via a linear program. Vague probabilities are produced as with Smith's theory if one assumes that a bettor is uncertain of the odds at which he/she is prepared to bet; or if one wishes different odds when betting for an event than when betting against it.

Suppes (1974) also developed an axiomatic system producing a form of vague probability. His work includes a combination of De Finetti's ideas and a finite version of Savage's structural axiom on infinite partitions. Domotor and Stelzer (1971) have performed some purely abstract work which gives results that may be interpreted similarly. The standard theories

of subjective probability assume DM's to be capable of comparing the desirability of any two decision alternatives and from this deduce the existence of an underlying subjective probability distribution. Rather than assume the existence of a total order among alternatives, one might assume only a semi-order. A semi-order may be derived from the concept of a "just noticeable difference," (jnd). A jnd has a technical definition, but for our purposes we may understand it by use of natural language. We assume that any two alternatives differing by at least a jnd can be distinguished as to desirability, whereas two alternatives not differing by a jnd will be judged to be of equal value. Assuming that such a quantity as the jnd exists, the alternatives are semi-ordered. It is shown that, in that case, the precise probabilities of Savage are replaced by vague probabilities. Note that as the jnd becomes arbitrarily small, then this system becomes equivalent to Savage's.

Each of the authors discussed above has thus developed axioms which will produce vague probabilities. Further, several of them have shown that, within the set of ranges provided by the vague probabilities, there exist numbers which satisfy the classical probability calculus. Smith (1961) refers to these as "medial odds;" Good (1962) refers to vague probabilities as meaning simply that there exist unknowable precise (classical) probabilities within the ranges indicated; and such values are easily deducible from the work of Domotor and Stelzer (1971) and of Suppes (1974). Further, as noted by Good in the discussion to Smith's (1961) paper, in order to provide a complete theory of rational behavior, medial utilities need also to be assumed. These values could again be viewed as lying between lower and upper utilities. Smith (1961) does indeed propose such a theory.

3.3 Vague Probabilities and Sensitivity Analysis

We would appear to have an axiomatic justification for the standard form of sensitivity analysis--the assessed "probabilities" and "utilities" fill the role of medial values, and the ranges are provided by the vague probabilities and utilities. However, these systems do not have the simple property discussed in Section 3.1, that lower and upper probabilities should each satisfy the probability calculus. In fact, equalities are replaced by inequalities. For example, introducing the notation P_* for lower probabilities, and P^* for upper probabilities, each axiomatic theory shows that for mutually exclusive events, A and B,

$$P_*(A \vee B) \geq P_*(A) + P_*(B), \tag{3.3.1}$$

and that

$$P^*(A \vee B) \leq P^*(A) + P^*(B). \tag{3.3.2}$$

In other words, it is not irrational in these systems for a DM to have narrower ranges for compound probabilities than might be deduced from the constituent vague probabilities. For example, setting $B = \sim A$,

$$P_*(A \vee \sim A) = P_*(X) = 1 = P^*(X),$$

so here the range is reduced to zero.

The difficulty here lies in the concept of the second-order interactions mentioned at the beginning of this section. If these are not modeled in some way, it is impossible to deduce the vague probability for $A \vee B$ from the component probabilities. In order to discover $P^*(A \vee B)$ and $P_*(A \vee B)$, the analyst must therefore ask the DM either to indicate how the imprecisions in the probabilities for A and B are linked, or else assess pro-

babilities for $A \vee B$ directly.

If such further information is not elicited, our best deduction for derived probabilities will be obtained by replacing the inequalities of 3.3.1 and 3.3.2 by equalities. By failing to make a greater effort in modeling the DM's belief structure, we have failed to capture it completely. Such a failure may not necessarily be bad, for time constraints in assessment may make this limited modeling effort desirable. Further, it may be unnecessary to achieve greater precision in the compound probabilities if it becomes apparent that this will not affect the recommended decision. However, using as a general strategy the technique of building a decision tree and then placing vague probabilities on the chance-nodes is inadequate. Yet this is precisely what is achieved by performing a standard sensitivity analysis upon a decision tree.

Once one has accepted that the point probabilities of the basic DA paradigm are insufficient for a full analysis, the basis for the "divide-and-conquer" strategy is removed. One can no longer concentrate on the component probabilities and assume the compound ones will take care of themselves. Rather, since one has acknowledged that a sensitivity analysis will be required, one should build it into the fabric of the analysis. Thus, since it is the compound probabilities which are of primary interest, the analyst's efforts should be directed towards these. The component vague probabilities should be assessed; the links in their imprecisions considered; and from this the composite vague probabilities deduced. These should also have been assessed directly and by comparing direct

and derived values the DM may be able to improve the assessments in an iterative process, and thus improve his/her coherence and, we hope, the decision. The point that is stressed is that the entire belief structure must be probed and modeled, and not just the "minimally-sufficient" component probabilities. When using vague probabilities in this way, we have a firm axiomatic foundation, based on any of the theories discussed in 3.2, for performing a rigorous sensitivity analysis.

If the above procedure is followed, the output of the DA will be ranges for the expected utility of each alternative. If no alternative dominates all the others, selection of the preferred alternative is not straight forward. We might use the medial values, as assessed in a standard DA, and use the ranges to indicate sensitivity. We might consider it appropriate to continue the assessment procedure in an attempt to narrow the ranges until there was no overlap. Alternatively, we have shown elsewhere (Freeling, 1980a) that a reasonable criterion for selection is to take that alternative X_i such that (with an obvious notation)

$$U^*(X_i) = \operatorname{Max}_j U^*(X_j)$$

and

$$U_*(X_i) = \operatorname{Max}_j U_*(X_j).$$

If no such alternative exists, then further elicitation is necessary. This procedure is a special case of the techniques using fuzzy probabilities, which are discussed in Section 3.5.

3.4 Criticisms of Vague Probabilities

Although the vague probabilities discussed above appear to answer several of the questions that we wished to address, we can see that there are still inadequacies. First, there is no formal calculus for incorporating the links in the imprecision. It would be desirable to be able to use the divide-and-conquer strategy, or at least to be able to guide an analyst in comparing assessments for holistic and decomposed probabilities. The procedure described in the previous section rests on intuition to make these comparisons.

Second, it may be argued that there is further information concerning the DM's belief structure that could be, yet is not, incorporated. This concerns the possible values of a probability within the indicated range. *Often a DM will feel that some values are more reasonable* (in some sense) than others; yet using vague probabilities this feeling cannot be considered.

A third consequence of the properties of vague probabilities concerns the representation of ignorance concerning an event. In classical probability the "Principle of Insufficient Reason" is usually invoked. The event space is partitioned into subsets which are assumed each to be equi-probable. The trouble with this is that the partition is usually arbitrary, and thus the probability induced for the event of interest may be a very poor representation of the state of belief (or ignorance). An appealing use of a vague probability is to say that ignorance concerning an event may be modeled by placing the lower

probability at zero, and the upper probability at one. This appears to capture the concept of ignorance by saying nothing about the likelihood of the event. However, as proven by Smith (1961), upper and lower probabilities as derived from vague probabilities satisfy Bayes' Theorem exactly. That is, although there are inequalities in Eqns. 3.3.1 and 3.3.2, in the corresponding equations generalizing Bayes' Theorem, equalities remain. Although this may appear a desirable consequence, it causes difficulties. For, when updating a prior vague probability [0, 1] on the receipt of new evidence using Bayes' Theorem, the resulting posterior upper and lower probabilities will remain unity and zero, respectively. This is a result of the well known fact that when using Bayes' Theorem, prior certainty cannot be shifted. The reason for this difficulty is apparent. Our initial assessment admits to the possibility that the event might be impossible or certain. Whatever subsequent evidence we obtain (apart from observation of the event itself) the theory forces us to harbor continuing suspicions about the certainty of the event or of its negation. We might attempt to circumscribe this problem by setting the probability for ignorance at $[\varepsilon, 1-\varepsilon]$ for small, positive ε. This would allow updating, but the choice of ε is critical, yet it appears to be arbitrary. Thus we are forced to conclude that vague probabilities provide little improvement in the modeling of ignorance.

A theory of belief based on upper and lower probabilities that addresses the first and third of these points is examined in Section 4.0. The second point is addressed in Section 3.5.

3.5 Second-Order and Fuzzy Probabilities

One of the weaknesses of vague probabilities, discussed in Section 3.4, is that there is no indication of whether different parts of the range of probabilities are (in some sense) considered more reasonable than others. Two theories of belief have been developed that attempt to cope with this problem. One (second-order probabilities) uses standard probability theory; the other (fuzzy probabilities) makes use of the new Theory of Fuzzy Sets (Zadeh, 1965).

The basic concept of second-order probabilities is simple. As proposed in Lindley, Tversky, and Brown (1979) and Tani (1978) one treats the imprecision in a probability assessment as a form of uncertainty and argues that this in turn should be modeled by the probability calculus. Thus, one builds a probability distribution over the probability. The method of Lindley, Tversky, and Brown postulates the existence of a "true" probability π which a DM attempts to access from his/her psyche but which, due to forms of measurement error, he/she can assess only as a value q, which is π together with some random error. In particular, in all calculations where π would normally be used, we use the continuous distribution $Pr(\pi|q)$. The expectation of this distribution may be used as the single value for π if such is deemed necessary.

This concept of second-order probabilities is not new. Savage (1954) was aware of the difficulty of assessing all probabilities with total precision, but he discarded the idea of hierarchical probabilities as being impracticable. I. J. Good has done a lot of work on this concept of "hierarchical" probabilities. He has recently written a review article of his own work

(Good, 1980) and we refer the reader to that for further discussion of this topic. One obvious difficulty is that the second-order assessments can also not be assessed precisely--indeed they may exhibit second-order incoherence. One might then make third-order assessments, but there is clearly an infinite regression possible here, and no indication that it will converge (although Good, 1980, argues that it must converge, else people would never come to any agreement concerning beliefs). Good (1962) uses this imprecision in second-order probabilities to explain the vagueness that must exist in the upper and lower probabilities of the previous section. In any case, the higher-order assessments become progressively less meaningful to a DM, and any simplicity that might be provided by a DA will be lost. A second difficulty pointed out by Savage (1954) is that the expectation of the distribution may fill the role that the first-order probability was considered unable to fill; i.e., it becomes a point estimate of the uncertainty. The effect of our acknowledgement of the imprecision in probability assessments is thus lost. Finally, an axiomatic foundation for these second-order probabilities would probably be unconvincing because the comparisons between events that form the basis of most axiomatizations of subjective probability would be far less intuitive when dealing with second-order events of the form "the probability of event A is x."

The general problem with these second-order probabilities, of which the above properties are merely symptoms, appears to be that we are now attempting to put too much structure upon the DM's imprecision

for an individual probability. As soon as assessed values are constrained to satisfy the probability calculus, a great deal is assumed about the ability to make judgments concerning these values. (The problems that we have observed with assumptions of this form are, of course, the motivation for all the work described in this paper.) It seems to the author that forcing the probability calculus upon these, necessarily vague, judgments of imprecision requires more complex effort of precisely the type we wish to replace! We are seeking a theory or technique that permits us to perform sensible types of sensitivity analysis upon basic probability assessments, and in order to do this, we wish to "separate out" our beliefs concerning likelihood of events from our imprecision and vagueness in those beliefs. We believe that this separation can be achieved by looking at the problem from the perspective of fuzzy set theory. We shall also attempt to show in this section that the fuzzy set theoretical concept is not as antithetical to the probabilistic view as is often suggested.

To provide for an easier exposition of these ideas, we shall look once again at second-order probabilities, and set up some notation. Suppose we are interested in two events A and B, and their probabilities p and q, respectively. Suppose further that A and B are mutually exclusive, so that $Pr(A \vee B) = p + q$. Then the second-order approach discussed above would be to assess probability distributions over p and q, and then (assuming non-interaction between these distributions, to be discussed later) to treat p and q as independent random variables

and to use the probability calculus to derive the probability distribution over their sum, r. As is well known, this sum has the density function calculated as the convolution of the densities of p and q. Specifically, if p has density

$f_p(x)$ $\quad$ $(x \varepsilon [0,1])$; and q has density

$f_q(y)$ $\quad$ $(y \varepsilon [0,1])$; then r has density

$f_r = f_p * f_q$, where * is defined by

$$f_r(z) = \int_{[0,1]} f_p(x) f_q(z-x)\,dx \qquad (3.5.1)$$

This may be rewritten as

$$f_r(z) = \int_{x+y=z} f_p(x) f_q(y)\,dx \qquad (3.5.2)$$

which equation we shall term the "Probabilistic Extension Principle," as this is the extension to probability distributions of the simple equation $p + q = r$.

Just as the density function, f, may be taken as the underlying aspect of the second-order probability, so the membership function, μ, is the basic concept of fuzzy probabilities. A fuzzy probability is represented by a function

$\mu_p(x)$ $\quad$ $(x \varepsilon [0,1])$, $\quad$ $\mu_p(x) \varepsilon [0,1]$.

We shall assume here that μ_p is continuous and well-defined. This

membership function may have any of several different interpretations. It has been interpreted variously as the possibility that x is the probability (Zadeh, 1978); the degree of truth of the statement "Pr(A) is x," (Watson, Weiss, and Donnell, 1979); the compatibility of the value x with the probability of A (Zadeh, 1977). A slightly different interpretation has been suggested by Freeling (1980a, c). His interpretation is related to the concept of vague probabilities discussed in the previous section. He uses the idea of a level set, which is defined as

$$P_a = \{x: \mu_P(x) \geq a\}, \quad a \varepsilon (0,1],$$

$$P_0 = \{x: \mu_P(x) > 0\}.$$

Clearly there is a one to one relationship between the set of all level sets, and the membership function. Then P_a is interpreted as the set of values for Pr(A) such that the degree of compatibility of each value with the probability is at least a. Note that P_a form a nested set of intervals;

$$\text{i.e.,} \quad a \leq b \rightarrow P_a \subseteq P_b.$$

The level set at level a is then the vague probability at a given level of confidence. So P_0 is the vague probability such that we are certain the range could be no broader, and P_1 is the vague probability which is the most restricted--we could not distinguish between the possibility of any such points. A fuzzy probability captures the idea that some values are seen as more possible than others. Our earlier papers (Freeling, 1980a, b) have discussed the mathematics of fuzzy probabilities. The reader is referred to those for further details, particularly with regard to assessing membership functions.

For our present purposes the important aspect of a fuzzy probability is that the imprecision in a probability assessment as modeled by a membership function is fuzzy, rather than probabilistic. This we understand as a conceptual, rather than an arithmetical, distinction. When dealing with an uncertain quantity, modeled by probability theory, we may use the expectation of a distribution as our best guess of that quantity. When dealing with an imprecise quantity, we assume that the quantity is inherently imprecise, or fuzzy. The fuzzy distribution is our "best guess"--no reduction of that can make sense.

This indeed was Zadeh's major motivation for the invention of Fuzzy Set Theory. He asserted that certain types of imprecision in human thinking cannot be appropriately modeled by probability theory. While this assertion remains untested and controversial, we believe that in our current context it provides the correct perspective. By modeling the imprecision as fuzzy, we avoid the trap discussed earlier of taking expected values to give point values for probabilities. Rather, we may continue in the spirit of our work on vague probabilities, and in the context of sensitivity analysis, to continue dealing with a range of probabilities.

An important question regarding the membership function is to ask to what calculus it should conform. There have been two suggestions for this which have achieved most attention. These are using either Max-min connectives, or product connectives. Again the reader is referred to the previous literature (Freeling, 1980a, c) for a discussion of the meaning of these terms. For this discussion, we may characterize

them in terms of the two corresponding "Fuzzy Extension Principles." Analogously to equation (3.5.2), if p and q are fuzzy probabilities, with membership functions $\mu_p(x)$ and $\mu_q(y)$, then $r = p + q$ may be defined as the fuzzy probability with membership function

$$\mu_r(z) = \max_{x+y=z} (\mathrm{Min}(\mu_p(x), \mu_q(y))), \tag{3.5.3}$$

if we use the Max-min connectives, and

$$\mu_r(z) = \int_{x+y=z} \mu_p(x)\mu_q(y)\,dx \tag{3.5.4}$$

if we use the product connectives.

As will be easily seen, we obtain two totally different theories, depending on whether we use equation (3.5.3) or equation (3.5.4). The Max-min connectives are the ones usually associated with fuzzy set theory. When interpreted in terms of level sets and degrees of confidence we believe equation (3.5.3) to be a good model. As discussed in an earlier paper (Freeling, 1980c), using the Max-min connectives means that each degree of confidence can be treated quite independently of each other. That is, if we are interested in the level set of Pr(A^B) at level a, then we need only known the level sets of Pr(A) and Pr(B) at level a. In other words, at any given level of confidence, fuzzy probabilities with Max-min connectives are simply vague probabilities (as defined in the previous section). Thus this theory answers the problem of vague probabilities that there is no indication of where in the range the probability is felt more certain to lie.

A theory of choice based upon the concept of fuzzy probabilities is developed in Watson, Weiss, and Donnell (1979), Freeling (1979, 1980a), and Adamo (1980). Each of these papers defines fuzzy utilities analogously to fuzzy probabilities, and Freeling looks in detail at various possible criteria for comparing fuzzy expected utilities. Dubois and Prade (1979) also address this issue of comparison. Although this theory may be viewed as a multi-level sensitivity analysis, similar arguments to those used in Section 3.3 show that it is an inadequate modeling of the situation to look at the fuzziness only in constituent elements of a decision tree, for such fuzziness may be linked. For example, if $Pr(A \wedge B)$ is required, and we have only the fuzzy probabilities for $Pr(A|B)$ and $Pr(B)$, then we have insufficient information. Instead, we should assess the imprecision in $Pr(A \wedge B)$ directly. Thus under this normative theory, we see once again that performing a sensitivity analysis subsequent to the main analysis is inadequate if the interaction between imprecision was previously unmodeled.

An axiomatic foundation for the Max-min fuzzy probabilities appears to be fairly easily derivable from the axioms for vague probabilities discussed in the previous section. A discussion of how this might be done has also been presented in our earlier papers. A small point that can be noted is that, when using the Max-min operations, one need not measure degrees of confidence on a continuous zero-one scale. Because the level sets are effectively disconnected, one may label them by qualitative factors; e.g., very confident, certain, etc. In this way we still perform a multi-level sensitivity analysis, but without demanding an excessive degree of extra information from the

decision maker. This idea has been explored by Whalen (1980). There remains, however, the problem of interaction. This problem can only be fully resolved by assessing all the fuzzy probabilities and utilities of interest.

If we were to use fuzzy probabilities with the product connectives, we would, of course, be using a different form of fuzzy sets. The reader will have noted that equation (3.5.4) is identical in format to equation (3.5.2). That is, the fuzzy probabilities are operated on in exactly the same way as second-order probabilities. The difference between (3.5.4) and (3.5.2) is purely a conceptual one: by treating the imprecision as fuzzy, and differentiated from probabilistic, we make explicit note of the differences between the two types of belief. As noted earlier, there is no concept of expectation in the fuzzy case; the probabilities, or expected utilities, should be left as fuzzy variables, rather than projected onto a single point estimate. Any attempt to quote single values would be without foundation. An axiomatic basis for such fuzzy sets may be hard to find, although the work of Hamacher (1976) may be of relevance. The general idea of these operators is to provide some compensation between high and low levels of confidence, thus narrowing the level sets. One could, therefore, view these operators as a surrogate way of accounting for the links in imprecision, since the effect of narrowing ranges is the same.

In conclusion then, we have shown that fuzzy probabilities, when interpreted in terms of level sets, are a natural extension of the concept of vague probabilities. They need not be viewed as an attempt to deny

the value of the concept of chance in decision making, as has been felt by some Bayesians, but rather as an attempt to extend the basic theory of probability so as to increase the applied usefulness of the theory. The point of our discussion of the formal equivalence of fuzzy probabilities and second-order probabilities, is to argue that when a distribution has been arrived at (such as $Pr(\pi|q)$ in Lindley et al., 1979) we should leave it at that, and not try to achieve further accuracy by looking at the expectation. This conclusion parallels that of Lindley (private communication) regarding the Lindley et al. work on RIJ.

3.6 Value of Coherence

A concept developed recently by the author (Freeling, 1980c, d) exploits the idea of fuzzy probabilities to derive a measure of the value of performing further analysis of a DM's belief structure. The concept is based on the raionale for considering extended theories of belief that we presented in Section 2.0. We take the view that with (hypothetical) perfect coherence a DM would be able to produce point probabilities, but due to imperfections the DM can produce only fuzzy probabilities. Performing a DA will reduce this potential incoherence: in the limit, to zero. The technique is similar to that used in Value of Information analyses. Prior to an analysis, we cannot know what the result of that analysis will be, but our initial assessments give us some information concerning the possible results. Specifically, the fuzzy membership functions indicate the possibility of various final results, and by a technique similar to that of "flipping the decision tree," the "value of perfect coherence" may be calculated.

Using the above observations as a basis, we may calculate the "value of perfect coherence." We note that such coherence will be of value only if we discover that the alternative we would have selected prior to the analysis was in fact suboptimal. If so, for a given set of coherent values, we may calculate the increased value of the improved decision over the prior one. Then the possibility of these being the coherent values is equal to the possibility of that being the value of coherence. In this way a possibility distribution over the value of coherence may be calculated. This may then be used to guide decisions concerning whether to pursue further analysis. The mathematics of this concept are discussed in Freeling (1980c, d).

This concept provides a powerful new tool for deciding the value of a DA. It fits in well with our perception of the role of decision analysis: namely reducing incoherence. It is not our intention to discuss the concept in detail here, but the following points should be noted.

(a) The value of increased but imperfect coherence may be similarly calculated.

(b) The concept is not confined to fuzzy probabilities. It may be similarly defined for second-order probabilities, in which case it is simply the value of information concerning our uncertainty in the probability. A vague probability may be viewed as simply a fuzzy probability with membership function unity over the range of the vague probability, and zero elsewhere. Then the value of coherence becomes an interval.

3.7 Conclusions

In this section we have discussed the various proposals that have been made to relax the standard assumptions of probability theory. We have shown each to be axiomatically justifiable, and of potential usefulness. We shall discuss our conclusions concerning the implications of these theories and of those discussed in the next two sections, in Section 6.0. For now we shall note that if faced with the choice of whether to use vague probabilities, or the richer fuzzy probabilities, we feel that this choice will depend on the context. Fuzzy probabilities require more effort in assessment, and experience has shown that the use of vague probabilities (in the form of sensitivity analysis) is often sufficient. Thus for a first pass at a problem, it is probably reasonable to use vague probabilities.

4.0 BELIEF FUNCTIONS

The extended theories of belief discussed so far in this section all have a common philosophical basis. They each take the DM's belief structure as fundamental. They do not take explicit account of the external stimuli that may have led to the DM adopting that belief structure. In this, the theories follow the theory of subjective probability, arguing that it is only the precision demanded by that theory which is unreasonable. Shafer (1976) with his theory of belief functions takes a different perspective. He views evidence as the fundamental concept. For this reason, his book is called, "A Mathematical Theory of Evidence," and he refers to his theory as an evidential theory of belief. This is not intended to mean that evidence should, in an objective manner, cause a DM to hold certain beliefs, but rather that the (subjective) beliefs held by the DM are the result of the DM's interpretation of the evidence presented to him/her.

The work presented by Shafer is an extension of previous work by Dempster (1967, 1968) on lower and upper probabilities as induced by multivalued mappings. Dempster placed his work directly in the context of lower and upper probabilities. Although Dempster's theory is contained in Shafer's, the latter downplays the role of lower and upper probabilities to concentrate on evidence. While that aspect of the theory is indeed the most original, several insights provided by Dempster's perspective,

which are of relevance to our present work, are omitted by Shafer. Although we shall present this section in Shafer's terminology, we shall point out where Dempster's perspective may be of use.

The theory differs from the previous ones in that it is developed from three fundamental axioms regarding the calculus of belief functions. It is not developed from any behavioral-type axioms, and does not attempt to describe the process of judgment by which the DM arrives at his belief function. This, as we shall see, greatly limits the possibility of applying these ideas directly to decision-aiding, for there is no indication of where the numbers should come from.

The theory becomes very complicated mathematically, and there is no possibility of a complete exposition in this brief overview. We shall, therefore, present the basics of the theory, and discuss it in the terminology of the previous sections. We shall attempt to show what behavioral assumptions and what sort of underlying philosophy would need to be accepted in order for this to be taken as the basis for a practical theory of belief. For an extremely lucid and complete exposition of Shafer's philosophy and mathematics, his book is unlikely to be bettered. Our interpretation of his work, as follows, is a personal one which should not be viewed as a précis of his ideas.

The inability of DMs to provide precise probabilities is not viewed as an imprecision arising from imperfect human information processing. Rather, except in rare circumstances, the amount of evidence available is considered insufficient to justify such precision. There is not,

as there is with the other extended theories, a difficulty in explaining why the numbers are imprecise--the DM's belief structure is in fact assumed to be precisely modeled by a belief function. In this way, the problem identified for the other theories of how to combine probabilities is not encountered. Because the belief structure is assumed to be precisely modeled, the calculus that is developed shows us exactly how the combination is to be effected.

The mathematical basis of Shafer's theory is similar to that of probability theory. The classical theory assumes that there is a probability "mass" of one which is distributed over the event space. In other words, the belief of the DM is viewed as a quantity which is apportioned over the possible events. Shafer similarly assumes belief to be quantifiable in terms of probability mass equal to unity. However, he argues that since one piece of evidence might support only a set of events, rather than a single event, the belief induced by that evidence should be apportioned to that set of events, and not to any particular event. For example, consider the annual Oxford v. Cambridge boat race (as have Smith, 1961; Brown and Lindley, 1981; Freeling, 1981a). An event of relevance to the outcome is a coin toss, for the winner of that can take an important inside bend on the River Thames. Our event space (in Shafer's terms, the "frame of discernment") is X = {WC, LC, WO, LO} where C means Cambridge wins the race, and W that they win the toss, and O and L are the negations of these events. So WO stands for the event that Cambridge wins the toss, but Oxford wins the race (at the time of writing a depressingly common event!). Then as we receive information that Cambridge has won the toss, but for some reason we

do not completely trust our source, we might assign probability mass 0.6 to the subset of $X = \{WC, WO\}$ which we shall call Y. The remaining mass of 0.4 is left unassigned to any particular subset but assigned to the full event space.

Formally, we look at the <u>power set</u> of X, 2^X, which is defined as the set of all subsets of X. Then we define a function

$$m: 2^X \to [0,1], \text{ the } \underline{\text{basic probability assignment}},$$

such that $m(\emptyset) = 0$ and $\sum_{A \subset X} m(A) = 1$,

In other words, m is a probability distribution over the power set of X. A belief function is defined as a function from $2^X \to [0,1]$ such that

$$\mathrm{Bel}(A) = \sum_{B \subset A} m(B), \tag{4.1}$$

for all $A \subset X$. Thus m(B) is the belief ascribed <u>precisely</u> to the subset B, and Bel(A) is the belief ascribed to A and to all subsets of A. The logic of this arises from the evidential nature of the theory. Any evidence that supports a subset B, must equally well support all subsets including B; for if B were to occur, $A \supset B$ would also occur.

To relate this to standard probability theory, recall that a probability distribution over X is a function

$$P: X \to [0,1]$$

such that $\sum_{x \in X} P(x) = 1$, and $P(A) = \sum_{x \in A} P(x)$ for all subsets $A \subset X$.

Now suppose that the basic probability assignment of a belief function were such that $m(B) = 0$ for all subsets of X with more than one event.

Then
$$\mathrm{Bel}(A) = \sum_{B \subset A} m(B) = \sum_{x \in A} m(x) \tag{4.2}$$

By comparison, we see that with such a basic probability assignment, the belief function is a probability function. The link with lower and upper probabilities becomes clearer if we define the function

$$P^*: 2^X \to [0,1] \text{ by } P^*(A) = 1-\mathrm{Bel}(\sim A) \text{ for all } A \subset X.$$

Then
$$\begin{aligned} P^*(A) &= \sum_{B \subset X} m(B) - \sum_{B \subset \sim A} m(B) \\ &= \sum_{B \subset X} m(B) - \sum_{B \cap A = \emptyset} m(B) \\ &= \sum_{B \cap A \neq 0} m(B) \geq \sum_{B \subset A} m(B) = \mathrm{Bel}(A) \end{aligned}$$

Interpreting this, Bel(A) is the total probability mass that we are certain lies within A, and is thus a lower bound on our possible belief in A; whereas P*(A) is the total probability mass on subsets that have some intersection with A. Such mass <u>might</u> therefore, support A, and indeed we see that P*(A) is an upper bound on our possible belief in A. For this reason, P*(A) is termed an upper probability, and Bel(A) may be viewed as a lower probability. Using the previous example, where $m(Y) = 0.6$ and $m(X) = 0.4$, we find that

$$\mathrm{Bel}(Y) = \sum_{B \subset Y} m(B) = m(Y) = 0.6$$

$$P^*(Y) = \sum_{B \cap Y \neq 0} m(B) = m(Y) + m(X) = 1$$

Thus the belief in Y must be at least 0.6, but it could be total. That is an indication of the fact that there is no evidence contradicting Y. Now the subset of X that is of interest to us is Z = {WC, WL}; i.e., Cambridge winning:

$$\mathrm{Bel}(Z) = \sum_{B \subset Z} m(B) = 0, \text{ since neither X nor Y are contained in Z.}$$

$$P^*(Z) = \sum_{B \cap Z \neq 0} m(B) = m(X) + m(Y) = 1, \text{ since } Y \cap Z = \{WC\} \neq \emptyset.$$

This is equivalent to total ignorance regarding the subset Z. This is reasonable since, without any further assumptions, the evidence on Y has told us nothing concerning Z. Note that this is in contrast with classical Bayesian theory, whereby we would be forced to appeal to some form of the "Principle of Insufficient Reason" to apportion out our probability mass over the singletons of X. Shafer argues that since we should be concerned solely with the evidence provided, such further assumptions are unjustified. Indeed, we may represent total ignorance (or lack of evidence) concerning X by taking the basic probability assignment $m(B) = 0$ for all $B \neq X$, and $m(X) = 1$. This is a mathematical expression of our not having any idea where our belief should be placed. Then it is easy to see that for all $B \neq X$, $\mathrm{Bel}(B) = 0$ and $P^*(B) = 1$, which, as discussed earlier, is our preferred expression of ignorance. Further, it is <u>not</u> the case with this theory that lower probabilities of 0 cannot be updated in the light of further evidence. In other words, belief functions do not necessarily satisfy Bayes' Theorem.

In order to understand heuristically why this is so, one needs to keep in mind the fundamental difference between vague probabilities and belief functions. As emphasized earlier, a vague probability is vague because of imprecision which no attempt is made to model. This imprecision is therefore carried through the analysis. Belief functions, however, produce lower and upper probabilities because the evidence, which is modeled precisely via the basic probability assignment, does not justify further precision. On receipt of further evidence, the belief structure may be altered so as to increase specificity and to narrow the range between lower and upper probabilities.

We have not so far defined the calculus which belief functions satisfy. It should be clearly understood that we do not refer here to calculating, for example, the degree of belief in (A and B) given the belief in A and B singly. Whereas with probabilities, and independence, we have $Pr(A \wedge B) = Pr(A)Pr(B)$, and similar expressions for vague or fuzzy probabilities; within the present theory such a question has no meaning. After the receipt of a given piece of evidence we model the entire belief structure, by assessing the basic probability assignment over every subset of X. Thus we assess $m(A)$, $m(B)$, and $m(A \wedge B)$, and then the upper and lower probabilities are already determined. The calculus we need to define concerns the rules for combining the belief induced by separate items of evidence. Shafer effects this via "Dempster's rule of combination." In order to explain this, it is best to take yet another perspective on the theory of belief functions.

The reader will have realized by now that the basis of the theory is the probability assignment, m, which is simply a probability distribution over the subsets of X. In fact, m defines a random subset of X. This is analogous to a random number. A random number is defined by a probability distribution over the real line, and may be interpreted as the result of a random experiment the result of which is a real number. A random set is also defined by a probability distribution, but over a set of subsets, and the realization is one of these subsets. So, for example, if $X = \{A, B\}$, then $2^X = \{0, \{A\}, \{B\}, \{A,B\}\}$, and the result of the experiment might be any of the four elements of 2^X. The same terminology applies to random subsets as to random numbers. In particular, we may talk of stochastic independence between random subsets--S_1 and S_2 are independent random subsets, over X, if and only if

$$\Pr\{(S_1, S_2) = (Y, Z)\} = \Pr(S_1 = Y)\ \Pr(S_2 = Z)$$

for all $Y, Z \subset X$.

Then a belief function may easily be related to the underlying random subset S, for

$$\mathrm{Bel}(A) = \sum_{B\subset A} m(B) = \sum_{B\subset A} \Pr(S=B) = \Pr(S\subset A).$$

Thus Bel(A) is the probability that the random subset is contained in A. Goodman (1980a, b) terms this the superset coverage function of S. Nguyen (1978) discusses the links between belief functions and random subsets, with regard to their abstract mathematical structure. He also shows that this work is interpretable in terms of the capacities of Choquet (1953-1954). If we have two pieces of evidence, giving rise

to belief functions Bel_1 and Bel_2 and to random subsets S_1 and S_2, the resultant belief function Bel is defined from its basic probability assignment

$$m(A) = \frac{Pr(S_1 \cap S_2 = A)}{Pr(S_1 \cap S_2 \neq \emptyset)}, \quad A \neq \emptyset \tag{4.3}$$

so that

$$Bel(A) = \frac{Pr(S_1 \cap S_2 \subset A)}{Pr(S_1 \cap S_2 \neq \emptyset)} .$$

Shafer does not provide an a priori argument for this definition, but relies on the reasonableness of the results developed from it. Dempster's original work (1967) in defining his rule of combination, justified it in a manner similar to the following heuristic argument. If we have two separate items of evidence, generating two random subsets, and we are interested in knowing the result of having both items simultaneously, let us consider the relevant experiments. We are seeking a random subset which would have the same underlying probability distribution as if we performed the two other experiments simultaneously. Were we to perform those experiments we should only observe those elements of X that were in the outcome of <u>both</u> experiments. Thus we would only observe $S_1 \cap S_2$. Therefore, m(A) should be $Pr(S_1 \cap S_2 = A)$, normalized by $Pr(S_1 \cap S_2 \neq 0)$ to ensure that the total probability mass is unity. (Recall that $m(\emptyset) = 0$ by definition of m.) Shafer does not in fact talk in terms of random sets, but he looks only at "entirely distinct bodies of evidence" which may thus be assumed to generate stochastically independent random subsets. Dempster's rule is thus stated in the form

$$m(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i)\, m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i)\, m_2(B_j)}$$

which i⌃ clearly equivalent to (4.3).

This ability to combine evidence is the cornerstone of Shafer's theory. To gain a better understanding, consider the following simple situation.

Suppose X to have just two subsets; A and ~A.
Then $2^X = \{\emptyset, A, \sim A, X\}$.
Suppose we have two distinct items of evidence, one supporting A, and the other ~A. Suppose also that the strength of each piece of evidence is such that

$$m_1(A) = 0.4 \quad \text{and} \quad m_2(\sim A) = 0.7,$$

with an obvious notation, that the remainder of the belief is assigned to X, so $m_1(X) = 0.6$, $m_2(X) = 0.3$. Then combining these via Dempster's rule we find that

$$m(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i)\, m_2(B_j)}{1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i)\, m_2(B_j)} = \frac{m_1(A)\, m_2(X)}{1 - m_1(A)\, m_2(\sim A)}$$

$$= .12/.72 = 1/6$$

Similarly,

$$m(\sim A) = m_1(X)\, m_2(\sim A)/0.72$$
$$= .42/.72 = 7/12,$$

and $m(X) = .18/.72 = 1/4.$

Thus $Bel_1(A) = 0.4$ and $P_1^*(A) = 1$

$Bel_2(A) = 0$ and $P_2^*(A) = 0.3$

and $Bel(A) = 1/6,\ P^*(A) = 5/12.$

These results appear to accord fairly well with our intuition. The conflicting pieces of evidence have reduced the degree of belief caused by each other, but the stronger evidence (the second) has had the greater impact. Consider now a case where each piece of evidence is itself conflicting. Let $m_1(A) = 0.4$ as before, but take $m_1(\sim A) = 0.6$. Similarly, let $m_2(\sim A) = 0.7$, but $m_2(A) = 0.3$. Then

$$m(A) = m_1(A)m_2(A)/(1 - [m_1(A)m_2(\sim A) + m_1(\sim A)m_2(A)])$$

$$= 0.12/0.54 = 2/9$$

$$m(\sim A) = m_1(\sim A)m_2(\sim A)/0.46$$

$$= 0.42/0.54 = 7/9. \qquad (4.4)$$

Note that here m_1 and m_2 both produce normal probability functions, for all the probability mass is on the singletons of X. So

$$Bel_1(A) = m_1(A) = 0.4, \text{ and } P_1^*(A) = 1 - 0.6 = 0.4.$$

Similarly

$$Bel_2(A) = 0.3 = P_2^*(A).$$

If we had written the effect of the first piece of evidence in the form $Pr_1(A) = 0.4,\ Pr_1(\sim A) = 0.6$

and of the second

$Pr_2(A) = 0.3$, $Pr_2(\sim A) = 0.7$, it would appear that we had a reconciliation problem. By pooling the two probability distributions (belief functions), we seem to have effected the reconciliation. Further, the pooled evidence as reflected in (4.4) is clearly also a probability distribution, so the theory appears sound.

The answer is $Pr(A) = Bel(A) = P^*(A) = 2/9$ and $Pr(\sim A) = 7/9$. Yet this result is very strange, for the reconciled value for Pr(A) is much less than both the original values. Such a reconciliation appears most unnatural. The reason for this difficulty lies at the root of the problem in applying Shafer's theory to real situations. It lies in a misinterpretation of the meaning of the numbers. Suppose the two probability distributions came from two different experts. Then the first expert might be saying that the likelihood of A occurring was the same as that of drawing a random number with last digit between zero and three inclusive. Similarly, the second expert would have chosen between zero and two inclusive. If we considered the experts to be independent and equally reliable, we would reasonably choose 0.35 as our best guess of the likelihood of A. When considering the two belief functions, however, a different perspective is relevant. Then the first expert is viewed as implying that there is some evidence supporting A, but more evidence refuting it. The second expert, independently, is providing evidence supporting A but more evidence refuting it. It is then reasonable to suppose that the two experts taken together provide even stronger evidence refuting A, and thus that our belief in A should be further

reduced, to below 0.3. The subtlety is that the numbers with belief functions do not signify likelihood, but rather strength of support. This distinction is crucial, for it means that even with Bayesian belief functions (Shafer's term for belief functions with the properties of probability distributions), there is no apparent way to assess the requisite numbers. When we talk about more general belief functions, it becomes even less clear what the numbers mean and how they should be assessed. Although the basic probability assignment has the formal properties of a probability distribution corresponding to a random subset, there is no apparent real-world experiment to which it could correspond. Therefore, it cannot be interpreted as a likelihood and it appears to be impossible to assess. Prade (1979) recognizes this, and states that m should not be termed a probability, but simply a "basic assignment." Shafer attempts to cope with the problem by talking about the "weight of evidence," and using this to define the corresponding degrees of belief, but he has to scale these weights arbitrarily, and the numbers once again lose their intuitive meaning. This difficulty of interpretation was already foreshadowed in the discussion to Dempster's (1968) paper; the contributions by Smith, Bartholomew and the response by Dempster each acknowledge his lower and upper probabilities to be different than those produced by a betting paradigm. This difference is further explored in the work by Cohen (1970, 1973, 1977, 1979, 1980) discussed in Section 5.0.

That this state of affairs obtains is unfortunate, for the notion of belief functions appears to answer many of the questions raised in earlier subsections. The entire belief structure is modeled, unlike

other theories that produce upper and lower probabilities, but we now see that, as with second-order probabilities, this is only achieved at the cost of requiring very difficult, if not impossible, assessments. The ability to pool different and possibly conflicting pieces of evidence would be an invaluable aid to decision making. Using probability theory, such combination is effected using Bayes' Theorem, which is only an efficient procedure when the evidence is presented in terms of events which are <u>known</u> to have occurred. (For a discussion of that, see Shafer, 1976.) To effect a "reconciliation" of the form discussed above is very difficult in the Bayesian framework--witness the literature produced on the subject, e.g., Morris (1974, 1977); Lindley, Tversky, and Brown (1979); Brown and Lindley (1981); and Freeling (1981). The entire concept of using evidence as the foundation for a *theory of belief* is very appealing, yet cannot be incorporated simply into the Bayesian framework.

For all these appealing features the price we have to pay is too great. Rather than relaxing the difficulty of the assessment task for the DM, we have made it more difficult by requiring a probability distribution over the power set of X rather than just X, and further there is no intuitive meaning to this distribution. One possibility to help alleviate this problem would be to use vague (or even fuzzy) probabilities for the basic probability assignment m, thus permitting a mix of modeled and unmodeled beliefs. We intend to look at this concept of "vague belief functions" in further research. We also note that Dempster's rule of combination applies only to independent pieces of evidence and often the links between pieces of evidence are all

too clear. The importance of this in the context of reconciliation is discussed by Freeling (1981). Shafer does not address this problem, but using the interpretation based on random subsets a possible extension becomes apparent. The links may be expressable in terms of non-independence between the subsets, and the pooling effected via equation (4.3). This concept, too, we shall be exploring in our further research.

For the present, however, we must reject the use of belief functions, while taking note of the perspective Shafer's theory gives us on how we should model belief. Simply put, the important points are:

1) the evidence impacting on belief is important, and
2) when taking into account vagueness in belief, the <u>entire</u> belief structure needs to be modeled, and not just a small ("minimally specified") part of it.

The way we can try and use these concepts in a standard decision analysis is discussed in Section 6.0.

5.0 ORDINAL MEASURES OF BELIEF

As discussed in Section 2.1, some researchers have looked at theories of belief from a perspective which is different from that of a range of probabilities. These workers appear to take issue not with the idea that belief in an event be representable by a single number, but rather they propose a different calculus to be appropriate when operating with these values. In fact, rather than use the product operators usually associate with probability theory, they advocate the use of either maximum or minimum. The theory of this type that is perhaps best known arises from Zadeh's theory of fuzzy sets (Zadeh, 1965) and in particular his theory of possibility (Zadeh, 1978). Two theories of similar form have been developed quite separately by L.J. Cohen (1970, 1973, 1977, 1979, 1980) who studies inductive (Baconian) probabilities; and by Shackle (1969) who defines degrees of surprise.

In this section we shall examine these theories. In particular we shall show that there is a very close link between possibility theory and inductive probability theory, and that each in turn may be viewed in the context of a restricted class of belief functions. In this way, we show that posssibility theory may be viewed as an evidentiary theory of belief; indeed possibilities may be considered as upper inductive probabilities.

Zadeh's Possibility Theory is a natural extension of his theory of fuzzy sets. Suppose we have a proposition "X is F", which is defined in terms of a membership function $\mu_F(x)$, for all $x \varepsilon X$. Then F gives us possibilistic information concerning the state of the world. Zadeh defines the possibility of x, $\Pi(x) = \mu_F(x)$. Then the possibility of $A \subset X$, is

$$\Pi(A) = \Pi(\bigcup_{x \varepsilon A} x) = \max_{x \varepsilon A} [\Pi(x)] = \max_{x \varepsilon A} [\mu_F(x)].$$

The basic rule which can be used to define this measure is that

$$\Pi(\emptyset) = 0, \ \Pi(X) = 1, \text{ and for all } A, B \subset X,$$

$$\Pi(A v B) = \mathrm{Max}(\Pi(A), \Pi(B)).$$

A great deal of literature has been devoted to the examination of the use of this maximum operator. We have reviewed and examined this literature at length in our previous work (Freeling, 1979, 1980a, c) and rather than discuss it here, we refer the interested reader to that work. Yager (1979) and Whalen (1980) have extended the ideas to produce a theory of choice. Each of these authors proceeds in a direction analogous to decision analysis by defining "fuzzy utilities" which combine with possibilities using Max-min connectives rather than addition and product. We have examined these theories in detail in Freeling (1980c) so here we shall simply comment that we do not believe the calculus to be sufficiently well-motivated to provide a convincing decision aid.

Cohen proposes his theory as a complete alternative basis for probabilistic reasoning to classical (Pascalian in his terminology) probability. He looks at inductive probabilities, using the notation that the probability of A is $P_I(A)$. Then the major rule of his theory is that

$$P_I(A \wedge B) = \text{Min}\ (P_I(A),\ P_I(B)). \qquad (5.1)$$

His theory is based on ideas of evidential support for a proposition, and a generalization of inductive logic. It would thus appear to have close links to the theory of belief functions. That this is indeed so we shall see shortly.

Shackle (1969) describes an entire theory of choice based on the degree of surprise. Such a degree of surprise is a function $S: 2^X \to [0, 1]$. It satisfies $S(A \vee B) = \text{Min}\ (S(A), S(B))$, and also that

$$\text{Min}\ (S(A),\ S(\sim A)) = 0.$$

Shackle's arguments for the above rules are based on an intuitive understanding of the way a degree of surprise should act. He develops a theory of choice based on these ideas, using the principle that the preferred alternative be the "most interesting." This theory follows in the spirit of mathematical economics by defining optimality in terms of contiguous tangent curves. Shackle has little to say about how the values in his theory are assessed.

Each of these measures of belief might seem highly antithetical to the theories discussed earlier in this section. However, in fact all three theories may be placed in the context of a restricted class of

belief functions. This fact was noted by Shafer (1976) in the case of Shackle's and Cohen's theories; and by Prade (1979) in the case of Shackle's and Zadeh's theories.

The class of belief functions that is relevant is that termed by Shafer consonant belief functions. These may be identified as belief functions such that the basic probability assignment, m, assigns non-zero mass only to elements in a nested chain; i.e., $\{A_i \subset X: m(A_i) > 0\}$ is of the form $A_1 \subset A_2 \subset A_3 \ldots \subset X$. Shafer terms such belief functions consonant because they betray no hint of conflict in the evidence. Specifically, the belief functions can never accord positive degrees of belief to both sides of a dichotomy. In fact, Shafer proves that

$$\mathrm{Min}(\mathrm{Bel}(A|B), \mathrm{Bel}(\sim A|B)) = 0 \text{ for all } A, B \subset X \text{ such that } \mathrm{Bel}(\sim B) < 1 \quad (5.2)$$

It can then be shown easily that in their formal properties, Cohen's inductive probability $P_I(A)$ corresponds to Bel(A); Shackle's degree of surprise S(A) corresponds to Bel(~A); Zadeh's possibility $\Pi(A)$ corresponds to P*(A).

For example, with this definition,

$$P_I(A \wedge B) = \mathrm{Bel}(A \wedge B) = \sum_{A_i \subset A \wedge B} m(A_i)$$

But $A_i \subset A \wedge B$ only if $A_i \subset A$ and $A_i \subset B$.

Choose k so that A_k is the largest A_i that is contained in $A \wedge B$. Then

$$\sum_{A_i \subset A \wedge B} m(A_i) = \sum_{i=1}^{k} m(A_i).$$

But if similarly A_n is the largest $A_i \subset A$; and A_m the largest $A_i \subset B$, we see that $A_i \subset A \wedge B$ if and only if $i \leq \mathrm{Min}(n,m)$. Therefore,

$$k = \mathrm{Min}(n, m),$$

so

$$\mathrm{Min}(P_I(A),\ P_I(B))$$

$$= \mathrm{Min}(\mathrm{Bel}(A),\ \mathrm{Bel}(B)) = \mathrm{Min}\Big(\sum_{i=1}^{n} m(A_i),\ \sum_{i=1}^{m} m(A_i)\Big)$$

$$= \sum_{i=1}^{k} m(A_i) = P_I(A \wedge B),$$ which is Equation 5.1 as required, Q.E.D.

(This is Shafer's Theorem 10.1 part (a).)

This perspective affords us interesting insights into the nature of the three theories. However, since Shackle's theory does not fit in with the main ideas of our work, and since we have not studied it in sufficient depth, we shall discuss it no further.

5.1 Inductive Probabilities

Cohen's work, as mentioned previously, arises out of similar considerations to Shafer's in that both are evidentiary theories of belief. Cohen is concerned with a generalization of inductive logic. Specifically, he looks at the situation when there is insufficient evidence to allow for certainty in an inference, but where nevertheless some inference can be made from the evidence.

The major difference between inductive probabilities and belief functions, and one which worries Shafer (1976, p. 224) is that inductive probabilities, by equation (5.1), appear to require consonance of evidence. In fact, as Cohen (1980) points out, such consonance is not strictly assumed. His theory does permit dissonance, but the way in which this is treated can not be interpreted in terms of belief functions. With consonance, Shafer and Cohen agree both in the formal properties of their theories, and also in the motivation behind their work. As Shafer suggests, such consonance may be appropriate for "inferential evidence," which can be interpreted as Cohen's "inductive reasoning," and also for some forms of statistical evidence.

However, in many real-life situations, dissonance in evidence is apparent. For example, judicial evidence will often provide support both for the hypothesis under consideration, and for its negation. It is with such evidence that Shafer and Cohen diverge. As previously discussed, Shafer continues to use belief functions to model the effect of such evidence. However, property (5.1) is lost. The effect of the contradictory evidence is to reduce the upper probability of the hypothesis. For example, suppose $X = \{A, \sim A\}$. Then if $m(A) = a > 0$, and $m(X) = 1-a$, the belief function is consonant. By (5.2), we see that since $Bel(A) = a > 0$, $Bel(\sim A)$ must equal zero, so $P^*(A)$ must be unity, as is easily checked. If instead, $m(\sim A) = b > 0$, and $a + b < 1$, then there is contradictory evidence, and the belief function is no longer consonant. Then although $Bel(A) = a > 0$, $P^*(A) = 1 - Bel(\sim A) = 1 - b < 1$. Thus

the effect of the contradictory evidence is to reduce the range of belief in A.

Cohen, however, retains property (5.1) when dissonant evidence is encountered. Thus if $P_I(A|E) = a$ and $P_I(\sim A|E) = b$, where E denotes the evidence received,

$$P_I(A \wedge \sim A|E) = \text{Min}(a, b) > 0; \text{ i.e., } P_I(\emptyset|E) > 0.$$

This situation is therefore not modeled simply by a belief function, since $P_I(\cdot\,|\,E)$ assigns positive belief to the empty set. Cohen accounts for this by taking $P_I(\sim E) = P_I(A \wedge \sim A|E)$. In other words, since E accords positive support to a contradiction, E must be either incomplete or mistaken. A consequence of this is that in Cohen's theory we have a parallel to (5.2), viz

$$\text{Min}(P_I(A|E), P_I(\sim A|E)) = 0 \text{ if } P_I(\sim E) = 0. \qquad (5.3)$$

As Cohen points outs, this equation embodies a generalization of proof by contradiction. When using that method of proof (also known as "reductio ad absurdum"), if we assume A and then arrive at a contradiction, we must conclude that A is false, provided we are satisfied with our rules of inference. Similarly, if we accept our evidence E in the present situation, and then arrive at positive degree of support for a contradiction $(A \wedge \sim A)$, we can then say that there is equal support for the falsity of E. Underlying these ideas is the concept that true evidence can point only to the truth, and cannot therefore be contradictory. Where such contradiction is exposed, our hypothesis should be altered in order to remove that contradiction. (One obvious

way of altering the hypothesis is to include in it the possibility of the evidence not being the truth, the whole truth, and nothing but the truth.)

A major advantage of inductive probability over belief functions arises out of the retention of property (5.2). Because the only operation performed on such probabilities is taking the minimum, we need only a finite number of "degrees of support." These may be made context-dependent, and defined in terms of the strength of support for the hypothesis under consideration. Then an hypothesis H may be considered proven "beyond a reasonable doubt" if $P_I(H|E)$ is greater than a given, pre-defined level of support. With conflicting evidence, H may be considered proven over H' due to "preponderance of evidence" if $P_I(H|E)$ is a given number of levels greater than $P_I(H'|E)$.

It should be emphasized that Cohen's ideas are not directly contradictory to those of standard subjective probability. He does not deny that Pascalian (chance) probabilities have a place in probabilistic reasoning, but rather that in certain inferential tasks inductive probabilities are more appropriate a model. Indeed, Cohen is stating explicity the difference discussed in the previous section between probabilities related to likelihood (Pascalian) and those related to evidential support (Baconian).

The confusion about the distinction between these two types of probability is apparent in the exchange between Cohen, and Kahneman and Tversky regarding the experimental work of the latter (Cohen, 1979,

1980; Kahneman and Tversky, 1979).

Cohen (1979) argues that many of the "fallacies" discovered by Kahneman and Tversky in human probabilistic reasoning (Kahneman and Tversky, 1973, 1974; Tversky and Kahneman, 1974) are in fact caused by their assuming Pascalian reasoning to be the appropriate normative model for analyzing their results, when in fact Baconian reasoning was. Kahneman and Tversky (1979) reply that Baconian probability does not correspond to our intuitive notion of chance, and is normatively unsound. Cohen in his rejoinder (Cohen, 1980) points out that the amount of inductively related evidence is the basis of his ideas; it is not chance.

Our suspicion is that in fact humans have both intuitive concepts, and that one of the difficulties is the semantic one that "probability" is being used to describe each concept. In some problems of inference the evidential idea may be used, whereas in others, and in choice problems, chance is used. If it is made unclear which concept we wish to test, some combination of the two might be used. Some evidence supporting this hypothesis is presented by Schum and Martin (1980), whereby neither theory of probability adequately described the experimental results. Rather, the experiments indicated probabilistic reasoning took a form intermediate between the two. It would be interesting to perform further experiments to see if one might separate out the two types of reasoning, by providing more explicit instructions to subjects. It would also be intriguing to develop a formal model of probabilistic reasoning which used both concepts simultaneously.

Of course, using inductive probabilities rather than belief functions to model evidence will mean that several capabilities are lost--in particular combination of distinct bodies of evidence appears to be a more subjective matter. It seems, though, that this may be reasonable, and that the strength of assumptions with belief functions simply is too much for practicability. Our current feeling, then, tends towards the use of inductive probabilities to model the impact of evidence, but further research is necessary to help us better understand the implications of each theory.

5.2 Possibility Theory

We may use the ideas discussed earlier in this section to provide a new perspective of the role of possibility theory as a theory of belief. We can see that possibility theory is complementary to Cohen's ideas, for defining

$$\Pi(A|E) = 1 - P_I(\sim A|E), \tag{5.4}$$

$$\begin{aligned}\Pi(A \vee B|E) &= 1 - P_I(\sim(A \vee B)\,|\,E) = 1 - P_I(\sim A \wedge \sim B|E) \\ &= 1 - \mathrm{Min}(P_I(\sim A|E),\ P_I(\sim B|E)) \\ &= \mathrm{Max}(1 - P_I(\sim A|E),\ 1 - P_I(\sim B|E)) \\ &= \mathrm{Max}(\Pi(A|E),\ \Pi(B|E)),\end{aligned}$$

as required for a possibility measure. These properties are of course just a formal equivalence. It seems reasonable, however, to extend the equivalence to the interpretation of the ideas, placing possibility theory in the status of an evidential theory of belief.

Although this status has not been explicity recognized before, so far as we are aware, such an interpretation of possibility theory is not inconsistent with previous work. Gaines (1976a, b) discusses how fuzzy set membership functions may be derived from a truth-functional multi-valued logic. This derivation has obvious links to Cohen's ideas of partial inductive proof. Proponents of fuzzy set theory have repeatedly denied that chance is the underlying concept of possibility (e.g., Zadeh, 1965, 1978). Further, the choice of the word possibility suggests the interpretation implied by (5.4). An event is possible to the extent that it is not disconfirmed by the evidence; i.e., to the degree that its negation is not confirmed.

By analogy with Shafer's definition of upper probability, we see that the possibility $\Pi(A|E)$ may be considered to be the upper inductive probability $P^*_I(A|E)$. Thus the unification of these two theories has been made possible. We anticipate that a more general theory of inference from limited evidence may be possible by considering both concepts, analogously to the way that upper and lower (chance) probabilities are proposed in Section 3.0. For example, proof by preponderance of evidence might be considered achieved if $P_I(H|E)$ were sufficiently high, and $\Pi(H'|E)$ sufficiently low. Also, this interpretation may help us in a reexamination of the theories of choice by Yager (1979) and Whalen (1980), and indicate in what contexts those theories may be reasonably applied.

Two further observations regarding the formal properties of possibility are appropriate here. First, we may borrow from Cohen's ideas that,

with conflicting evidence, $\Pi(X|E)$ need not be unity. Then

$$\Pi(A \vee {\sim}A|E) = \mathrm{Max}[\Pi(A|E), \Pi({\sim}A|E)] = a < 1.$$

Then we may conclude

$$\Pi(E) = a < 1.$$

This parallels Cohen's definition that $P_I({\sim}E) = 1 - a$.

Second, there is a possible confusion between two distinct concepts arising from fuzzy set theory. As discussed in Section 3.0, a fuzzy set A is usually described by a membership function, $\mu_A(x)$, and the intersection $A \wedge B$, the union $A \vee B$, and the complement ${\sim}A$, defined by

$$\mu_{A \wedge B}(x) = \mathrm{Min}(\mu_A(x), \mu_B(x)), \tag{5.5}$$

$$\mu_{A \vee B}(x) = \mathrm{Max}(\mu_A(x), \mu_B(x)), \tag{5.6}$$

$$\mu_{\sim A}(x) = 1 - \mu_A(x). \tag{5.7}$$

Possibility measures are defined via

$$\Pi(A \vee B) = \mathrm{Max}(\Pi(A), \Pi(B)). \tag{5.8}$$

The similarity between equations (5.6) and (5.8) is striking, and might tempt one into further extending the definition of a possibility measure by

$$\Pi(A \wedge B) = \mathrm{Min}(\Pi(A), \Pi(B)) \quad \text{and/or} \tag{5.9}$$

$$\Pi({\sim}A) = 1 - \Pi(A). \tag{5.10}$$

Each of these would, however, be a mistake. Note first that if we assume (5.10) we may derive (5.9), for

$$
\begin{aligned}
\Pi(A \wedge B) &= \Pi(\sim(\sim A \vee \sim B)) \\
&= 1 - \Pi(\sim A \vee \sim B) && \text{by 5.10} \\
&= 1 - \mathrm{Max}(\Pi(\sim A),\ \Pi(\sim B)) && \text{by 5.8} \\
&= 1 - \mathrm{Max}(1-\Pi(A),\ 1 - \Pi(B)) && \text{by 5.10} \\
&= \mathrm{Min}(\Pi(A),\ \Pi(B)).
\end{aligned}
$$

But if both 5.8 and 5.9 are true, then

$$\Pi(X) = \Pi(A \vee \sim A) \quad = \quad \mathrm{Max}[\Pi(A),\ \Pi(\sim A)]$$

and

$$\Pi(\emptyset) = \Pi(A \wedge \sim A) = \mathrm{Min}[\Pi(A),\ \Pi(\sim A)].$$

Thus for normal situations, one of $\Pi(A)$ and $\Pi(\sim A)$ must be unity, and the other zero. This, however, holds true for all A, so the possibility measure is no more than a binary, Boolean measure.

5.3 Summary and Conclusions

In this section we have shown that Zadeh's possibility theory and Cohen's theory of inductive probability are closely related, and that each is a theory of belief on evidence rather than chance. We note that both the evidential concept and the chance concept may be intuitive to humans. We also suggest that confusion of the term "probability" may have led to inappropriate use of each concept, both in decision making and in laboratory experiments and thus, to inconsistency. Insufficient work has been performed to permit any strong conclusions to be drawn; nor have we yet developed any practicable decision-or-judgment-aids from these ideas. However, we are convinced that further work

on these measures is justified of both experimental and methodological type. We are hopeful that this will soon lead to an improved range of applied decision and inference aids.

6.0 IMPLICATIONS FOR DECISION ANALYSIS

In the previous sections we have examined the various theories of belief, and looked at their potential value in decision aiding. We have not, in the scope of the work reported here, been able to take the next step and incorporate these ideas into producing an improved methodology. In this section, however, we present a tentative outline of the kind of changes that are indicated and the shape that such a methodology might take, based upon the present work.

6.1 Divide and Conquer?

The concept of "divide-and-conquer" has been the foundation stone of much practical decision analysis. The idea is that a DM will find it easier to make judgments concerning complicated matters if the problem is decomposed into its constituent parts. The DM is thus required to make more, but we hope simpler, judgments. As a basic concept we believe this to be sound, but that it is carried too far in many decision analyses. As will be understood from the discussion of the extended theories, the divide and conquer strategy takes no account of the links between the imprecision in the various assessments. We thus run into difficulties during the sensitivity analysis if we examine only the sensitivities to the assessed values, for it is at this level that the links in imprecision are of paramount importance.

However, this is typically the way in which a sensitivity analysis is conducted. A major reason for this would appear to be the status that the decision tree has enjoyed within practical DA. As von Winterfeldt (1980) points out:

> "trees so much dominate decision analytic structures that structuring is often considered synonymous to building a tree."

It is usually tacitly assumed that once the decision tree has been built (which, to be sure, will be the result of an iterative process), the structuring is complete. Then "all" that remains will be eliciting the requisite probabilities and utilities and exploring the numerical results. However, we consider it vital that the sensitivity analysis be considered during the structuring process. The decision tree structure is without doubt adequate for the basic analysis as a vehicle for using the divide-and-conquer strategy. It is for precisely that reason, however, that the decision tree is inadequate as the basis for a sensitivity analysis: dividing will not conquer. As stressed throughout the paper, the whole belief structure needs to be explored. This can be achieved by recognizing the fact throughout the analysis, and assessing imprecise probabilities and utilities; probing for inconsistency and links in imprecision all along. The decision tree may still be retained for its simplicity and clarity to the DM, but more assessments than the minimally specified set indicated by the tree should be elicited.

Because we are dealing throughout with imprecise probabilities and

utilities, there is always the possibility of increasing the precision. The value of doing this can be calculated using the concept of the Value of Coherence (Section 3.6) at each stage of the assessment procedure.

A second difficulty with using a decision tree with a few specified probabilities is that the nature and role of the evidence leading to probability assessments is completely obscured. While we do not go so far as do Shafer and Cohen in claiming that the probability calculus is inappropriate for modeling the effect of evidence on belief, we do feel that the nature and amount of evidence should be shown explicitly. Especially when DA is to be used for multiple DMs, or as part of a public decision making process, it will often be important to attempt to show <u>why</u> one assessor feels a certain value (or range of values) should be assigned to a given probability. This would reduce disagreement concerning such values, or at least help pinpoint from whence such disagreement arises.

We hope in our continuing research to provide a means of quantifying the weight of evidence, based on inductive probability or belief functions. For now, however, we are constrained to using chance-tested probabilities in a decision aid. The impinging evidence should be described qualitatively, perhaps with the aid of an influence diagram (Howard and Matheson, 1980) or of a hierarchical inference structure (Kelly and Barclay, 1973; Martin, 1980; Schum and Martin, 1981).

6.2 Suggested Methodology

The procedure that we advocate for performing a decision analysis is thus required to satisfy the following four criteria:

a) It should deal throughout with some form of imprecise probabilities and utilities;

b) The entire belief structure of the DM should be explored, rather than just the target probabilities;

c) The role of evidence in finding the assessed values for probabilities should be made explicit; and

d) The value of performing further exploration of the DM's belief structure should be exhibited.

We tentatively put forward the following procedure for performing such a decision analysis. We expect those details which are at present left vague will become clearer after we have used these ideas in some practical analyses.

Prior to a detailed formal analysis, a quick pre-modeling of the problem should be performed, with the general types of available options specified. Their values may then be assessed in the form of (very) imprecise expected utilities, and a very rough value of coherence analysis performed. Assuming that the value of coherence appears to be sufficient to justify a full analysis, the following steps should be performed.

Stage 1. Model and structure the problem as with standard DA. Build the decision tree, without placing values on any of the unknown variables. As at present, the analyst must be prepared to update the model throughout the analysis if this appears to be necessary.

Stage 2. List the uncertain events which are of relevance to the problem. This should be a full list; so, for example, if $Pr(A \wedge B)$ is needed in the problem, the list should include $Pr(A)$, $Pr(B)$, $Pr(A \wedge B)$, $Pr(A/B)$, $Pr(B/A)$.

Stage 3. Discover, together with the DM, what items of evidence impinge upon the uncertain events in question. Show these links, possibly in diagrammatic form.

Stage 4. Start to quantify the uncertainties in the events. This quantification may be in terms of any of the theories of imprecise probabilities discussed earlier. We feel that in most situations vague probabilities will suffice. The links in imprecision should be taken into account by direct assessment of all the probabilities listed in Stage 2. At this stage inconsistencies should be pointed out and reconciliations performed to help improve the assessments. "Best guess" medial probabilities, taking the place of ordinary precise probabilities, may also be assessed if the DM feels comfortable with this.

Stage 5. Assess the imprecise utilities in a manner analogous

to the quantification of uncertainty.

<u>Stage 6</u>. Using the appropriate calculus, compute the vague or fuzzy expected utilities. Use the best guess values, if available, as in a standard DA, and use the vagueness to look at sensitivity to the results.

<u>Stage 7</u>. Point out to the DM any inconsistencies the foregoing may have raised. Look at the value of coherence to help decide whether further analysis and specification are necessary. Iterate by returning to Stage 4 if necessary.

<u>Stage 8</u>. Present the results of the analysis to others with an interest. Try and pinpoint where disagreements arise--in the modeling; in the evidence considered relevant; in the impact that evidence is considered to have. Use the analysis as a basis for discussion to help reduce differences. Remember, the DA should be seen as a <u>guide</u>, not an oracle.

There remain of course many gaps in the above algorithm. These will best be filled in after the methods have been used in some case studies, which is the necessary next stage of research. The steps presented do not represent a radical departure from present-day decision analysis. We do not feel that there is a need for such a change, since the current procedures are usually effective. The difference is primarily one of emphasis: by emphasizing the entire belief structure; the importance of relevant evidence; and the

fundamental importance of sensitivity analysis and the decision maker's participation in it the analyst will be able to use the tools of decision analysis more effectively.

7.0 REFERENCES

Adamo, J.M. Fuzzy decision trees. Fuzzy Sets and Systems, 1980, 4, 207-219.

Boole, G. An investigation of the laws of thought. New York: Reprinted by Dover (1958), 1854.

Brown, R.V., and Lindley, D.V. Improving judgment by reconciling incoherence. To appear in Theory & Decision, 1981.

Cohen, L.J. The implications of induction. London: Methuen, 1970.

Cohen, L.J. A note on inductive logic. The Journal of Philosophy, 1973, LXX(2), 27-40.

Cohen, L.J. The probable and the provable. Oxford: Clarendon Press, 1977.

Cohen, L.J. On the psychology of prediction: Whose is the fallacy? Cognition, 1979, 7, 385-407.

Cohen, L.J. Whose is the fallacy? A rejoinder to Daniel Kahneman and Amos Tversky. Cognition, 1980, 8, 89-92.

Choquet, G. Theory of capacities. Annals of the Institut Fourier, (Universite Grenoble) 1953-1954, 5, 131-296.

de Finetti, B. Theory of probability. London: Wiley, 1974.

Dempster, A.P. Upper and lower probabilities induced by a multivalued mapping. Annals of Mathematical Statistics, 1967, 38, 325-339.

Dempster, A.P. A generalization of Bayesian inference. Journal of the Royal Statistical Society, Series B, 1968, 30, 205-232.

Domotor, Z., and Stelzer, J. Representation of finitely additive semiordered qualitative probability structures. Journal of Mathematical Psychology, 1971, 8, 145-158.

Dubois, D., and Prade, H. Decision making under fuzziness. In Gupta, M.M., Ragade, R.K., and Yager, R.R. (Eds.), Advances in Fuzzy Set Theory and Applications, 1979.

Edwards, W. The theory of decision making. Psychological Bulletin, 1954, 51, 380-417.

Edwards, W., and Tversky, A. (Eds.). Decision-making: Selected readings. Middlesex, England: Penguin Books, Ltd., 1967.

Freeling, A.N.S. Decision analysis and fuzzy sets. Unpublished masters dissertation, Cambridge University, Engineering Department, 1979.

Freeling, A.N.S. Fuzzy sets and decision analysis. IEEE Transactions on Systems, Man, and Cybernetics, 1980, SMC-10(7). (a)

Freeling, A.N.S. _Alternative frameworks for the reconciliation of probability assessments_ (Technical Report 80-4). Falls Church, VA: Decision Science Consortium, Inc., November, 1980. (b)

Freeling, A.N.S. Possibilities versus fuzzy probabilities--Two alternative decision aids. Submitted in 1980 for inclusion in H.J. Zimmerman et al. (Eds.), _Decision analysis through fuzzy sets_. TIMS/ORSA Studies in Management Science, 1980. (c)

Freeling, A.N.S. _Fuzzy probabilities and the value of coherence_. Paper presented at the International Congress on Applied Systems Research and Cybernetics, Acapulco, Mexico, December 12-16, 1980. (d)

Freeling, A.N.S. Reconciliation of multiple probability assessments. _Organizational Behavior and Human Performance_, 1981, _28_(3), in press. (a)

Freeling, A.N.S. _Reconciliation of inconsistent ratio judgments_ (Working Draft). Falls Church, VA: Decision Science Consortium, Inc., 1981. (b)

French, S. _On the methodology of decision analysis I: What is a normative theory of decision?_ Manchester, England: University of Manchester, Department of Decision Theory, March, 1979. (a)

French, S. _On the methodology of decision analysis II: Some thoughts on Bayesian decision analysis_. Manchester, England: University of Manchester, Department of Decision Theory, June, 1979. (b)

Gaines, B.R. Foundations of fuzzy reasoning. _International Journal of Man-Machine Studies_, 1976, _8_, 623-668. (a)

Gaines, B.R. Fuzzy reasoning and the logics of uncertainty. _Proceedings of the 6th International Symposium on Multiple Valued Logic, IEEE, 76CH1111-4C_, 1976, 179-188. (b)

Goldsmith, R.W. Studies of a model for evaluating judicial evidence. _Acta Psychologica_, 1980, _45_, 211-221.

Good, I.J. Rational decision. _Journal of the Royal Statistical Society, Series B_, 1952, _14_, 107-114.

Good, I.J. Subjective probability as the measure of a non-measurable set. In Nagel, E., Suppes, P., and Tarski, A. (Eds.), _Logic, Methodology and Philosophy of Science_. Stanford, CA: Stanford University Press, 1962.

Good, I.J. Some history of the hierarchical Bayesian methodology. In J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.F.M. Smith (Eds.), _Bayesian Statistics, Proceedings of the 1st International Meeting, Valencia, Spain, May 28-June 2, 1979_. Valencia, Spain: University Press, 1980, 489-519.

Goodman, I.R. _Identification of fuzzy sets with a class of canonically induced random sets_. Paper presented at the 19th IEEE Conference on Decision and Control, Albuquerque, New Mexico, December, 1980. (a)

Goodman, I.R. Fuzzy sets as random level sets: Implications and extensions of the basic results. Paper presented at the International Congress on Applied Systems Research and Cybernetics, Acapulco, Mexico, December 11-16, 1980. (b)

Hamacher, H. On logical connectives of fuzzy statements and their affiliated truth values. Paper presented at the Third European Meeting on Cybernetic Systems Research, Vienna, 1976.

Howard, R.A., and Matheson, J.E. Influence diagrams. Menlo Park, CA: SRI International, January, 1980.

Jungermann, H. Structural modeling of decision problems. Paper presented at the 88th Annual Convention of the American Psychological Association, Montreal, September 1-5, 1980.

Kahneman, D., and Tversky, A. On the psychology of prediction. Psychological Review, 1973, 80, 237-251.

Kahneman, D., and Tversky, A. Subjective probability: A judgment of representativeness. In Stael von Holstein, C.A.S. (Ed.), The Concept of Probability in Psychological Experiments, 1974, 25-48.

Kahneman, D., and Tversky, A. On the interpretation of intuitive probability: A reply to Jonathan Cohen. Cognition, 1979, 7, 409-411.

Kahneman, D., and Tversky, A. Prospect theory: An analysis of decision under risk. Econometrica, March 1979, 47(2), 263-291.

Keeney, R.L., and Raiffa, H. Decisions with multiple objectives: Preferences and value tradeoffs. New York: Wiley, 1976.

Kelly, C.W., III, and Barclay, S. A general Bayesian model for hierarchical inference. Organizational Behavior and Human Performance, 1973, 3(10), 388-403.

Koopman, B.O. The axioms and algebra of intuitive probability. Annals of Mathematics, 1940, 41, 269-292. (a)

Koopman, B.O. The bases of probability. Bulletin of the American Mathematical Society, 1940, 46, 763-774. (b)

Lindley, D.V., Tversky, A., and Brown, R.V. On the reconciliation of probability assessments. Journal of the Royal Statistical Society, Series A, 1979, 142(2), 146-180.

Martin, A.W. A general algorithm for determining likelihood ratios in cascaded inference (Report No. 80-03). Houston, TX: Rice University, Department of Psychology, November, 1980.

Morris, P.A. Decision analysis expert use. Management Science, 1974, 20, 1233-1241.

Morris, P.A. Combining expert judgments: A Bayesian approach. Management Science, 1977, 23(7), 679-693.

Nau, R.F. Coherent assessment of subjective probability. Berkeley, CA: University of California, Operations Research Center, March, 1981.

Nguyen, H.T. On random sets and belief functions. Journal of Mathematical Analysis and Applications, 1978, 65, 531-542.

Phillips, L.D., and Edwards, W. Conservatism in a simple probability inference task. Journal of Experimental Psychology, 1966, 72, 346-354.

Prade, H.M. Nomenclature of fuzzy measures. Seminar on theory of fuzzy sets, Johannes Kepler Universitat, Linz, Austria, September 24-29, 1979.

Raiffa, H. Decision analysis. Reading, MA: Addison-Wesley, 1968.

Savage, L.J. The foundations of statistics. New York: Wiley, 1954.

Schum, D.A., and Martin, A.W. Probabilistic opinion revision on the basis of evidence at trial: A Baconian or a Pascalian process? (Report No. 80-02). Houston, TX: Rice University, Department of Psychology, October, 1980.

Schum, D.A., and Martin, A.W. Assessing the probative value of evidence in various inference structures (Report No. 81-02). Houston, TX: Rice University, Department of Psychology, March, 1981.

Shackle, G.L.S. Decision, order, and time in human affairs (2nd ed.). Cambridge, 1969.

Shafer, G. A mathematical theory of evidence. Princeton, NJ: Princeton University Press, 1976.

Slovic, P. From Shakespeare to Simon: Speculations-and some evidence-about man's ability to process information. Oregon Research Institute Monograph, 1972, 12(12).

Smith, C.A.B. Consistency in statistical inference and decision (with discussion). Journal of the Royal Statistical Society, Series B, 1961, 23, 1-25.

Smith, C.A.B. Personal probability and statistical applications. Journal of the Royal Statistical Society, Series A, 1965, 128, 469-499.

Suppes, P. The measurement of belief. Journal of the Royal Statistical Society, Series B, 1974, 36, 160-175.

Tani, S.N. A perspective on modeling in decision analysis. Management Science, 1978, 24, 1500-1506.

Tversky, A., and Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 1974, 185, 1124-1131.

von Winterfeldt, D. Structuring decision problems for decision analysis. Acta Psychologica, 1980, 45, 71-93.

Wallsten, T.S. Measurement and interpretation of beliefs. In Jungermann, H., and de Zeeuv, G. (Eds.), Decision making and change in human affairs, 1977, 369-393.

Watson, S.R., Weiss, J.J., and Donnell, M.L. Fuzzy decision analysis. IEEE Transactions on Systems, Man, and Cybernetics, 1979, SMC-9, 1-9.

Whalen, T. Risk minimization using L-fuzzy sets. Proceedings of the International Conference on Cybernetics and Society, IEEE 80CH1555-2, 1980, 821-825.

Yager, R.R. Possibilistic decisions. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9, 388-342.

Zadeh, L.A. Fuzzy sets. Information and Control, 1965, 8, 338-353.

Zadeh, L.A. Linguistic characterization of preference relations as a basis for choice in social systems. Erkennis, 1977, 11, 383-410.

Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1978, 1, 3-28.

DISTRIBUTION LIST

Office of Naval Research

Stuart L. Brodsky
Group Leader/Mathematics
Code 411MA
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Neal D. Glassman
Code 411MA
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Charles J. Holland
Code 411MA
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

CDR Kent Hull
Code 410B
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Mr. Gerald Malecki
Engineering Psychology Programs
Code 442
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Charles R. Paoletti
Administrative Contracting Officer
Room 718
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Dr. James C. T. Pool
Division Leader/Mathematical
& Physical Sciences
Code 411
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

J. Randolph Simpson (10 copies)
Scientific Officer, Code 431
Naval Analysis Programs
Mathematical and Information
Sciences Division
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Dr. Martin A. Tolcott
Division Leader, Psychological
Sciences Division
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Dr. Thomas C. Varley
Group Leader
Code 411OR
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Edward J. Wegman
Group Leader, Code 411S&P
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217

Office of Naval Research
Eastern/Central Regional Office
666 Summer St.
Boston, MA 02210

DISTRIBUTION LIST (continued)

Other Government Agencies

Defense Technical Information
Center (12 copies)
Cameron Station
Building 5
Alexandria, VA 22314

Naval Research Laboratory (6 copies)
Attn: Code 2627
Washington, D.C. 20375

Supplemental Distribution

Professor Kenneth J. Arrow
Department of Economics
Stanford University
Stanford, CA 94305

Michael Athens
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

Dr. David Castanon
Alpha-Tech
Burlington, MA 01803

Dr. Ward Edwards
Director
Social Sciences Research Institute
1042 W. 36th Street
Denney Research Building
Los Angeles, CA 90007

Supplemental Distribution (cont.)

Dr. Peter Farquhar II
Graduate School of Administration
Voorhies Hall
University of California
Davis, CA 95616

Irwin R. Goodman
Naval Ocean Systems Center
Code 7232
San Diego, CA 92152

Professor James March
Graduate School of Business
Administration
Stanford University
Stanford, CA 94305

Dr. Amos Tversky
Department of Psychology
Stanford University
Stanford, CA 94305

Professor Robert Wilson
Department of Business Administration
Stanford University
Stanford, CA 94305

Professor L. Zadeh
Department of Engineering
University of California
Berkeley, CA 94701